ISSN 0956-8549-703

Estimating the Quadratic Covariation Matrix for an Asynchronously Observed Continuous Time Signal Masked by Additive Noise

By

Sujin Park
Oliver Linton

FINANCIAL MARKETS GROUP
DISCUSSION PAPER 703

April 2012

Sujin Park has a BSc Econometrics and Mathematical Economics, MSc Statistics and PhD Statistics from London School of Economics. Her research interests are inference for continuous time processes and asymptotic theory. Oliver Linton is Professor of Political Economy at the University of Cambridge. Previously, Professor of Econometrics at the London School of Economics and Research Fellow at the Financial Markets Group. Formerly Professor of Economics at Yale University. Fellow of the Econometric Society, Fellow of the Institute of Mathematical Statistics and Fellow of the British Academy. He is the author of over one hundred articles in the areas of econometric theory and financial econometrics. Within financial econometrics he is interested in high frequency data and in low frequency panel data issues. **Any opinions expressed here are those of the authors and not necessarily those of the FMG. The research findings reported in this paper are the result of the independent research of the authors and do not necessarily reflect the views of the LSE.**

Estimating the Quadratic Covariation Matrix for an Asynchronously Observed Continuous Time Signal Masked by Additive Noise *

Sujin Park
London School of Economics
London WC2A 2AE, United Kingdom
Oliver Linton
University of Cambridge
Cambridge,CB3 9DD, United Kingdom

January 15, 2012

*Sujin Park is a Doctoral Candidate, Department of Statistics, London School of Economics, Houghton Street, WC2A 2AE, United Kingdom (email: s.park4@lse.ac.uk). Oliver Linton is a Professor and Chair in Political Economy, Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD, United Kingdom (email: obl20@cam.ac.uk). We would like to thank the European Research Council for financial support. We thank Per Mykland for his invaluable comments and suggestions and Victor Todorov for his feedback.

Abstract

We propose a new estimator of multivariate ex-post volatility that is robust to microstructure noise and asynchronous data timing. The method is based on Fourier domain techniques, which have been widely used in discrete time series analysis. The advantage of this method is that it does not require an explicit time alignment, unlike existing methods in the literature. We derive the large sample properties of our estimator under general assumptions allowing for the number of sample points for different assets to be of different order of magnitude. The by-product of our Fourier domain based estimator is that we have a consistent estimator of the instantaneous co-volatility even under the presence of microstructure noise. We show in extensive simulations that our method outperforms the time domain estimator especially when two assets are traded very asynchronously and with different liquidity and when estimating the high dimensional integrated covariance matrix.

Keywords: Quadratic covariation, Fourier transform, Long run variance estimator, Market microstructure noise

1. INTRODUCTION

There have been many advances in the theory and application of volatility measurement from high frequency data. The ex-post measure of volatility called the quadratic variation has been the focus of much attention. The theory has been developed in a series of papers including: Andersen, Bollerslev, Diebold and Labys (2001), Barndorff-Nielsen and Shephard (2002, 2004) and Mykland and Zhang (2006). This work has been recently extended to take account of what is called microstructure noise when an underlying efficient price diffusion is distorted by measurement error in papers by: Zhang, Mykland, and Aït-Sahalia (2005), Zhang (2006), Kalnina and Linton (2008), Aït-Sahalia, Mykland and Zhang (2010b), Barndorff-Nielsen, Hansen, Lunde and Shephard (2008, 2011), and Jacod, Li, Mykland, Podolskij and Vetter (2009). In the multivariate case an additional issue arises, namely that the observations are asynchronous, i.e., transactions occur at different time points for different assets. Hayashi and Yoshida (2005) proposed an estimator of the integrated covariance that does not require synchronization. However, their estimator is inconsistent under the presence of microstructure noise. Malliavin and Mancino (2009) proposed a Fourier domain approach that does not require data alignment but they did not work out the theoretical results when the noise is present. Estimators addressing both the non-synchronicity and the microstructure noise have been proposed by Zhang (2010), Barndorff-Nielsen, Hansen, Lunde and Shephard (2011) and Aït-Sahalia, Fan and Xiu (2010a). The estimators are consistent with convergence rates respectively $O_p(n^{1/6})$, $O_p(n^{1/5})$ and $O_p(n^{1/4})$. The first two papers require aligning the data, although the consistency of their estimator is robust to the alignment. However, the hidden cost of data alignment and non-synchronicity for these estimators are that the sample size n that appears in the convergence rate is the sample size of aligned data. Also, the drawback of Zhang (2010) and Aït-Sahalia et al. (2010a) is that the estimator cannot be generalized to dimensions higher than two unless the covariance matrix is estimated element-wise, which in turn does not guarantee the positive definite estimator. See Park and Linton (2011) for a more detailed survey.

The goal of this paper is to propose a new estimator of the general multivariate volatility measure that is robust to microstructure noise and to asynchronous data timing. The method is based on Fourier domain techniques, which have been widely used in discrete time series to estimate the long run variance. The key advantage of this method is that it does not require an explicit time alignment. Our results allow for the unbalanced case where one series has many more observations than another, which is quite common in intra-day financial time series. In Section 2 we give a set up of the model and assumptions regarding the sampling scheme. In Section 3, we propose a Fourier domain based estimator of integrated covariance Section 4 presents the asymptotic properties of the proposed estimator without and with the presence of microstructure noise. The Fourier method is further extended to estimate the instantaneous covariance matrix and some economically interesting scalar functions of the integrated covariance matrix. We carried out extensive simulations and the empirical analysis reported in Section 5.

A word on notation. For scalars a and b, $a \wedge b$ and $a \vee b$ denote the minimum and maximum value. For a series $t_{i,j}$, denote $\Delta t_{i,j} = t_{i,j} - t_{i-1,j}$, and for any function g, let $\Delta g(t_{i,j}) = g(t_{i,j}) - g(t_{i-1,j})$. We use $\longrightarrow_p$ to denote convergence in probability, and $\Longrightarrow$ to mean stable convergence described in the Appendix. For real sequences a_n and b_n, $a_n \simeq b_n$ means $a_n = b_n + o_p(b_n)$. For a matrix $\mathbf{A}$, $\|\mathbf{A}\|_2 = \mathrm{tr}(\mathbf{A}^\intercal \mathbf{A})^{1/2}$. Let L denote the discrete time lag operator, so that $LX_t = X_{t-1}$.

2. THE MODEL AND ASSUMPTIONS

2.1 Efficient Price and Parameter of Interest

The following assumption describes the general setting used throughout the paper.

Assumption 1. *The efficient price process follows a Brownian semimartingale. For a $d \times 1$ vector of logarithmic prices $\mathbf{P}(t) = [P_1(t), \ldots, P_d(t)]^\intercal$ defined on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, \mathbb{P})$, we have*

$$\mathbf{P}(t) = \int_0^t \boldsymbol{\mu}(u)du + \int_0^t \boldsymbol{\sigma}(u)d\mathbf{W}(u),$$

where $\boldsymbol{\mu}(u) = [\mu_1(u), \ldots, \mu_d(u)]^\intercal$ *is a vector of predictable locally bounded drifts and* $\boldsymbol{\sigma}(u)$ *is a symmetric* $d \times d$ *matrix of locally bounded cádlág processes with* $\int_0^t \boldsymbol{\sigma}(u)\boldsymbol{\sigma}(u)^\intercal \otimes \boldsymbol{\sigma}(u)\boldsymbol{\sigma}(u)^\intercal du < \infty$ *a.s.* $\mathbf{W}(u)$ *is a* $d \times 1$ *vector of independent Brownian motion and is independent from the volatility process.*

The matrix $\int_0^t \boldsymbol{\sigma}(u)\boldsymbol{\sigma}(u)^\intercal \otimes \boldsymbol{\sigma}(u)\boldsymbol{\sigma}(u)^\intercal du$, which we call integrated quarticity, appears in the asymptotic variance of the estimator discussed later. The assumption of locally bounded drift and diffusion coefficient are required to apply Girsanov's theorem to remove the drift term in the theoretical derivation. Consider the discrete time grid $0 = t_0 < \cdots < t_n = T$, where T is fixed, and let $\mathbf{P}(t_i)$ denote the log price at those points. The quadratic covariation matrix of $\mathbf{P}$ over a time interval $[0, t], t \leq T$ is defined by

$$[\mathbf{P}, \mathbf{P}]_t = \operatorname*{plim}_{n \to \infty} \sum_{i; t_i \leq t} \{\mathbf{P}(t_i) - \mathbf{P}(t_{i-1})\}\{\mathbf{P}(t_i) - \mathbf{P}(t_{i-1})\}^\intercal, \tag{1}$$

where the limit is finite and well defined with probability one. Under Assumption 1, this is almost surely equal to the integrated covariance matrix

$$[\mathbf{P}, \mathbf{P}]_t = \int_0^t \boldsymbol{\sigma}(u)\boldsymbol{\sigma}(u)^\intercal du. \tag{2}$$

A natural estimator of (2) is the finite sum given in the definition of quadratic variation, which is called the Realized Covariance. Barndorff-Nielsen and Shephard (2002) showed that the Realized Covariance is unbiased and is a $\sqrt{n}$ consistent estimator of the integrated covariance under Assumption 1 and assuming synchronous trading. Throughout this paper we will reserve the square bracket to denote the quadratic variation, following the convention in the stochastic processes literature. The objective of this paper is to consistently estimate the integrated covariation matrix (2). The integrated covariance is related to the covariance matrix of prices by

$$\operatorname{cov}\{\mathbf{P}(t)\} = E\{\int_0^t \boldsymbol{\sigma}(u) d\mathbf{W}(u) (\int_0^t \boldsymbol{\sigma}(u) d\mathbf{W}(u))^\intercal\} = \int_0^t E\{\boldsymbol{\sigma}(u)\boldsymbol{\sigma}(u)^\intercal\} du = E[\mathbf{P}, \mathbf{P}]_t,$$

where the second equality follows from Itô's formula. Let $[\mathbf{P}, \mathbf{P}] := [\mathbf{P}, \mathbf{P}]_T$. We will denote the (i,j)- th element of an instantaneous covariance matrix by $\Sigma_{i,j}(u) = \{\boldsymbol{\sigma}(u)\boldsymbol{\sigma}(u)^{\intercal}\}_{i,j}$. The j-th diagonal element gives an integrated variance $[P_j, P_j] = \int_0^T \Sigma_{j,j}(u)du$.

Two problems are present in estimating (2). First, prices of different assets are observed at different times. Second, observed prices are distorted by noise and do not satisfy Assumption 1. We propose below an estimator that is robust to these two problems. We will examine in detail the two problems in the following sections.

2.2 Sampling scheme

In this section we describe the main assumptions we make on the observation times. We allow for unequally spaced and asynchronous observation times.

Assumption 2. *The time span is fixed and scaled to vary between $[0, 2\pi]$. We observe log prices at discrete time points: $0 = t_{0,\ell} < \cdots < t_{n_\ell,\ell} = 2\pi$ for $\ell = 1, \ldots, d$, where n_ℓ is the total number of observations for the ℓ-th asset. The discrete time points are allowed to be stochastic and assumed to be independent of the price and volatility process. The total number of observation points n_ℓ is large and $n := \min_\ell(n_\ell) \to \infty$. Unless otherwise stated, all convergence below holds with probability one. For all $a, b, \ell \in \{1, \ldots, d\}$:*

1. *The discrete time points satisfy $\sup_{0 \le i < n_\ell}(t_{i,\ell} - t_{i-1,\ell}) = O(n_\ell^{-1})$.*

2. *Denote the interval $I_{i,a} = [t_{i-1,a}, t_{i,a})$ and $I_{j,b} := [t_{j-1,b}, t_{j,b})$. Define the empirical quadratic covariation of time by*

$$\mathcal{Q}^{(n)}_{abcd}(t) = (n_a \wedge n_b \wedge n_c \wedge n_d) \sum_{i,j,k,l: t_{i,a}, t_{j,b}, t_{k,c}, t_{l,d} < t} (t_{i,a} \wedge t_{j,b} - t_{i-1,a} \vee t_{j-1,b})(t_{k,c} \wedge t_{l,d} - t_{k-1,c} \vee t_{l-1,d}) \times 1_{\{I_{i,a} \cap I_{j,b} \neq \emptyset\}} 1_{\{I_{k,c} \cap I_{l,d} \neq \emptyset\}} 1_{\{I_{min\{(i,a),(j,b)\}} \cap I_{min\{(k,c),(l,d)\}} \neq \emptyset\}}, \quad (3)$$

where $I_{min\{(i,a),(j,b)\}}$ denote $I_{i,a}$ if $n_a < n_b$ and $I_{j,b}$ otherwise. $I_{min\{(k,c),(l,d)\}}$ is equivalently defined. The empirical quadratic covariation satisfies $\mathcal{Q}^{(n)}_{abcd}(t) \longrightarrow \mathcal{Q}_{abcd}(t)$ as $n_a \wedge n_b \wedge n_c \wedge n_d \to \infty$, where $\mathcal{Q}_{abcd}(t)$ is continuously differentiable.

3. *The degree of non-synchronicity satisfies* $\sup_{i,j} |t_{i,a} - t_{j,b}| 1_{\{I_{i,a} \cap I_{j,b} \neq \emptyset\}} = O((n_a \wedge n_b)^{-1})$. *Given any set of* $\{t_{i,a}, t_{j,b}\}$ *such that* $n_a < n_b$, *we assume that*

$$\sup_{0 \leq j \leq n_b} \#\{t_{j,b} \in [t_{i-1,a}, t_{i,a}) | 1_{\{I_{i,a} \cap I_{j,b} \neq \emptyset\}}\} = O\left(\frac{n_a \vee n_b}{n_a \wedge n_b}\right).$$

The expression in Assumption 2.2 specializes to $\mathcal{Q}_{aa}^{(n)}(t) = n_a \sum_{i:t_{i,a}<t} (\Delta t_{i,a})^2$, which appears in the asymptotic variance of the integrated variance estimator and $\mathcal{Q}_{aabb}^{(n)}(t) = (n_a \wedge n_b) \sum_{i,j:t_{i,a},t_{j,b}<t} \Delta t_{i,a} \Delta t_{j,b} 1_{\{I_{i,a} \cap I_{j,b} \neq \emptyset\}}$, which appears in the asymptotic variance of the integrated covariance estimator. The expression (3) appears in the asymptotic covariance between the integrated covariance estimators. The Assumption 2 does not restrict the ratio of sample sizes of different assets to be bounded away from zero or infinity. One asset can be allowed to be much more liquid than the other. If Assumption 2.1 is further restricted to $\inf_i \Delta t_{i,\ell} = O(n_\ell{}^{-1})$ and $\sup_i \Delta t_{i,\ell} = O(n_\ell{}^{-1})$, then Assumption 2.3 is implied. Define $\{T_l^{(ab)}\}_{1 \leq l \leq N_T^{(ab)}} := \{t_{i,a} \cup t_{l,b}, i = 1, \ldots, n_a, l = 1, \ldots, n_b\}$, where $N_T^{(ab)}$ is a total number of data points for union of time stamps. The sample size of the union of time stamps, $N_T^{(ab)}$ is of order $O(n_a \vee n_b)$. We introduce here some notation we will use in the sequel. Denote the average interval size for asset ℓ by $\overline{\Delta t}_\ell := 2\pi/n_\ell$. When comparing asset a and asset b, denote for convenience the average interval size of the more liquid asset by $\widetilde{\Delta t}_{ab} = 2\pi/(n_a \vee n_b)$. We may drop the asset index whenever it is obvious.

3. ESTIMATION

3.1 Our Estimator

We propose to use the Fourier domain approach, which does not require data alignment at all. A nonparametric method based on the Fourier analysis of returns was first introduced by Malliavin and Mancino (2009). Frequency domain techniques are widely used in discrete time series analysis. One important and related application is the estimation of the long-run variance of a stationary time series analysis, which is equal to the spectral density at a frequency zero. We draw a natural link of such traditional method to the estimation of the quadratic covariation of a continuous time processes.

The Fourier basis given by $\{g_t(q) := \exp(iqt), q \in \mathbb{Z}\}$ where $i = \sqrt{-1}$ and $\overline{g_t(q)}$ denoting its complex conjugate, constitutes an orthonormal basis on the interval $t \in [0, 2\pi]$,

$$\frac{1}{2\pi}\int_0^{2\pi} g_t(k)\overline{g_t(j)}dt = \begin{cases} 1 & \text{if } k = j \\ 0 & \text{otherwise.} \end{cases}$$

We can express the continuous time signal $\{\mathbf{\Sigma}(t)\}_{t\in[0,2\pi]}$ as a linear combination of the Fourier basis with coefficient denoted by $\mathcal{F}(\mathbf{\Sigma})(q)$ for $q \in \mathbb{Z}$

$$\Sigma(t) = \frac{1}{2\pi}\sum_{q=-\infty}^{\infty} \mathcal{F}(\mathbf{\Sigma})(q)\exp(iqt), \tag{4}$$

and its Fourier pair by

$$\mathcal{F}(\mathbf{\Sigma})(q) := \int_0^{2\pi} \exp(-iqt)\mathbf{\Sigma}(t)dt, \qquad q = 0, \pm 1, \pm 2, \dots. \tag{5}$$

This is the continuous time Fourier transform of an instantaneous covariation matrix and at $q = 0$ we have the integrated covariance. We will propose an estimator for the above general form in (5). The above Fourier pair suggests that once we estimate the Fourier coefficient by $\widehat{\mathcal{F}}(\mathbf{\Sigma})(q)$, we may reconstruct the signal by replacing the infinite sum by the finite sum

$$\widehat{\mathbf{\Sigma}}(t) = \frac{1}{2\pi}\sum_{q=-n}^{n} \widehat{\mathcal{F}}(\mathbf{\Sigma})(q)\exp(iqt).$$

By Assumption 1, we have $\{\mathbf{\Sigma}(t)\} \in L^2([0, 2\pi])$ which guarantees that (4) is finite and $\|\widehat{\mathbf{\Sigma}}(t) - \mathbf{\Sigma}(t)\|_2 \to_p 0$. We next show how we can estimate (5) from the Fourier transform of the return process. We define the continuous time Fourier transform of return $dP_\ell(t)\,, \ell = 1, \dots, d$ satisfying Assumption 1

$$\mathcal{F}(P_\ell)(\alpha) = \int_0^{2\pi} \exp(-i\alpha t)dP_\ell(t), \qquad \alpha = 0, \pm 1, \pm 2, \dots \tag{6}$$

where the integral is a stochastic integral. The discrete Fourier transform of the ℓ-th asset is

$$\mathcal{F}_n(P_\ell)(\alpha) = \sum_{j=1}^{n_\ell} \exp(-i\alpha t_{j,\ell})\Delta P_\ell(t_{j,\ell}). \tag{7}$$

Let $\mathcal{F}_n(\mathbf{P})(\alpha) = \{\mathcal{F}_n(P_1)(\alpha), \ldots, \mathcal{F}_n(P_d)(\alpha)\}^\intercal$ for $\alpha \in \mathbb{Z}$ denote the vector of such Fourier transforms. Denote the amplitude window by $K_H(\cdot) : [-\pi, \pi] \to \mathbb{R}$. It is defined by $K_H(\lambda) := HK(H\lambda)$, where the spectral window $K(\cdot)$ satisfies the following assumption.

Assumption 3 *The spectral window $K(\lambda), \lambda \in [-\pi, \pi]$ satisfies the following conditions:* (i) $\int_{-\pi}^{\pi} K(\lambda)d\lambda = 1$, $\int_{-\pi}^{\pi} \lambda K(\lambda)d\lambda = 0$; (ii) $\int_{-\pi}^{\pi} |K(\lambda)|^2 d\lambda < \infty$, $\int_{-\pi}^{\pi} |\lambda K(\lambda)|^2 d\lambda < \infty$ *and* $\int_{-\pi}^{\pi} |\lambda^2 K(\lambda)|^2 d\lambda < \infty$*;* (iii) $K(\lambda) \geq 0, \forall \lambda \in [-\pi, \pi]$.

Our proposed estimator of (5) is given by

$$\widehat{\mathcal{F}}(\mathbf{\Sigma})(q) = (\widehat{\mathcal{F}}(\Sigma_{i,j})(q))_{i,j} = \sum_{|\alpha| \leq m/2} K_H(\lambda_\alpha)\mathcal{F}_n(\mathbf{P})(\alpha)\mathcal{F}_n(\mathbf{P})(q-\alpha)^\intercal, \tag{8}$$

where for $\rho(n) := \max_{\ell=1,\cdots,d} n_\ell$, we define $\lambda_\alpha = 2\pi\alpha/\rho(n)$, for $\alpha \in \mathbb{Z}$. We let $m = \rho(n)/H$ where the bandwidth $H \to \infty$ and $\rho(n), m \to \infty$ as $n \to \infty$. We are smoothing λ_α over the interval $[-\pi/H, \pi/H]$ where H controls the width of the smoothing window. We name our estimator, *Fourier Realized Kernel.* For $q = 0$, we may define the realized cross periodogram between assets 1 and 2 by $I_{12}(\alpha) := \mathcal{F}_n(P_1)(\alpha)\mathcal{F}_n(P_2)(-\alpha)$. Then $(1,2)$-th element of $\widehat{\mathcal{F}}(\Sigma)(0)$ is given by kernel smoothing the realized cross periodogram around the zero frequency

$$\widehat{\mathcal{F}}(\Sigma_{1,2})(0) = \sum_{|\alpha| \leq m/2} K_H(\lambda_\alpha) I_{12}(\alpha). \tag{9}$$

The condition (iii) in Assumption 3 guarantees that the estimators defined in (9) is p.s.d.

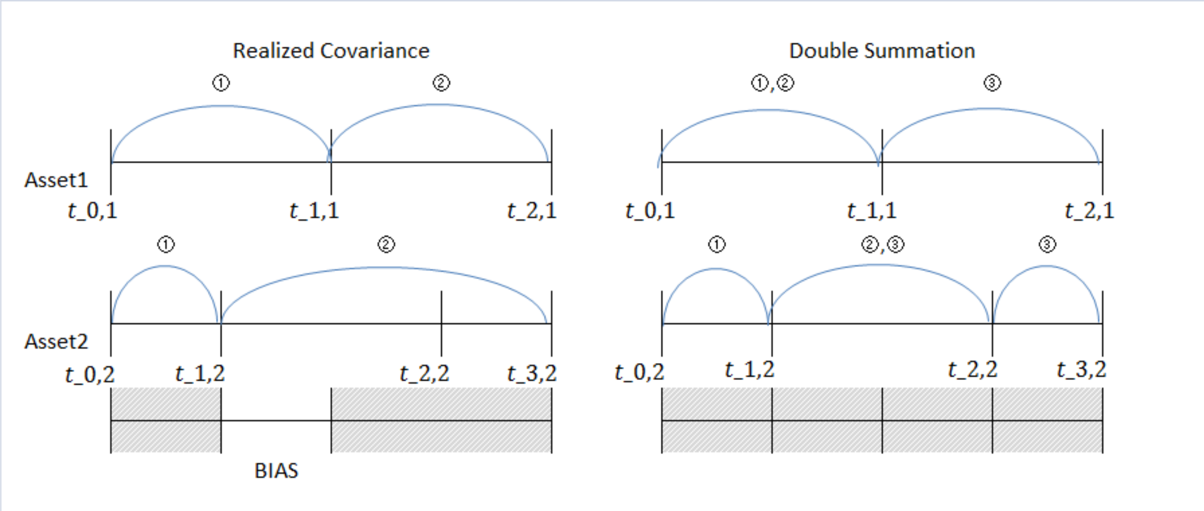
3.2 Comparison with some Time domain estimators

We first make a comparison with the covariation estimator of Hayashi and Yoshida (2005). Their estimator is a realized cross periodogram at zero frequency over the interval that overlaps i.e $HY = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \Delta P_1(t_{i,1})\Delta P_2(t_{j,2}) 1_{\{I_{i,1}\cap J_{j,2}\neq\emptyset\}}$. Then the centered realized cross periodogram can be decomposed into $I_{1,2}(0) - \int_0^{2\pi} \Sigma_{12}(t)dt = M_1 + M_2$, where

$$M_1 = HY - \int_0^{2\pi} \Sigma_{12}(t)dt \quad M_2 = \sum_{i,j} \Delta P_1(t_i)\Delta P_2(s_j) 1_{\{I_{i,1}\cap J_{j,2}=\emptyset\}}.$$

Hayashi and Yoshida (2008) showed that $\sqrt{n}M_1$ is asymptotically zero mean Gaussian, i.e., HY is unbiased estimator and achieves $\sqrt{n}$ consistency when microstructure noise is not

Figure 1: Toy Example [a]



[a]NOTE: Suppose $P_1(t) = P_2(t) = B(t)$, an independent Brownian motion. Then $[P_1, P_2](1) = \int_0^1 dt = 1$. Let $P_1(t_{i,1})$ be observed at $\{0, 1/2, 1\}$ and $P_2(t_{i,2})$ at $\{0, 1/4, 3/4, 1\}$. The refresh time grid is the same as the time stamp of the first asset in this case. The double summation estimator $\sum_{i,j} \Delta P_1(t_{i,1})\Delta P_2(t_{j,2})$ does not have a bias induced by aligning the non-synchronously observed data.

present. The realized periodogram is unbiased but inconsistent due to the extra term in M_2 which has a zero mean but it is a leading order term of $O_p(1)$. Figure 1 clarifies how the realized periodogram at zero frequency does not suffer from the bias due to non syncronicity of observation points. For data that is synchronized at $\{\tau_i\}$, we may define a realized autocovariance function $\gamma_{12}(h) = \sum_i \Delta P_1(\tau_i)\Delta P_2(\tau_{i-h}), h \in \mathbb{Z}$, where $\sum_i = \sum_{h<i\leq n}$ for $h \geq 0$, and $\sum_i = \sum_{1\leq i\leq n+h}$ for $h < 0$. Given a smoothing window in time domain $k(\cdot)$, the Realized Kernel proposed by Barndorff-Nielsen et al. (2011) is

$$\sum_{|h|<n} k\left(\frac{h}{H}\right) r_{12}(h) = \sum_{i,j=1}^{N} \Delta P_1(\tau_i)\Delta P_2(\tau_j) k\left(\frac{i-j}{H}\right). \tag{10}$$

We recognize that the estimator in (9) can be expressed in a similar form

$$\sum_{|\alpha|\leq m/2} K_H(\lambda_\alpha) I_{12}(\alpha) = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \Delta P_1(t_i)\Delta P_2(s_j) k_H(t_i - s_j), \tag{11}$$

where we defined the scaled bandwidth $\hbar = H\widetilde{\Delta t}$, in which case $k_H(t_i - s_j) := k(\frac{t_i - s_j}{\hbar})$. The lag window $k(\cdot)$ and spectral window $K(\cdot)$ are fourier transform pairs and using the Parseval's identity, the Assumption 3 can be translated to condition on the lag windows.

When time stamps are equally spaced and synchronous, i.e. $\tau_i = \tau_j + (i-j)2\pi/n$, the realized cross periodogram is a fourier transform of the realized autocovariance, i.e., $I_{12}(\alpha) = \sum_{h=-n+1}^{n-1} e^{-i\alpha h 2\pi/n}\gamma_{12}(h)$. Then our estimator (9) can be expressed as smoothing the realized autocovariances. When data is not syncronous, using all the data and implement (11) delivers superior estimator, which we will show in the next section. We also note that the relation between the smoothed periodogram to estimate the spectrum and data tapering (i.e., Fourier transforming the weighted return) is analogous to the relation between our estimator and the pre-averaging estimator of Christensen, Kinnebrock and Podolskij (2010).

4. ASYMPTOTIC PROPERTIES

4.1 Without Microstructure Noise

We consider the case where the sample sizes of different assets may not be of the same order of magnitude. We require the following rate condition.

ASSUMPTION 4. *H is a bandwidth satisfying $H \propto n^{\alpha}$ with $\alpha \in (0,1)$ so that we have as $n \to \infty$, $H \to \infty$ and $n/H \to \infty$. Also assume that $(n_a \vee n_b)/(n_a \wedge n_b) = o(H)$ for all $a, b \in \{1, \ldots, d\}$.*

REMARK Let $\beta \geq 1$ be a degree of liquidity parameter so that $n_a \vee n_b = O((n_a \wedge n_b)^{\beta})$. Then, Assumption 4 implies that $1 \leq \beta < 2$ for all $a, b \in \{1, \ldots, d\}$. By balancing the squared bias and the variance given in Appendix, the optimal bandwidth is given by $H = C_0 n^{\alpha^*}$, $\alpha^* = (4\beta - 3)/5$, where $C_0 \in (0, \infty)$. Then the convergence rate of the estimator under the optimal bandwidth is given by $(n_1 \wedge n_2)^{\vartheta}$, $0 < \vartheta = (4 - 2\beta)/5 \leq 2/5$. The result makes intuitive sense that for unbalanced sample sizes, the estimator converges at slower rate than the balanced case, $n^{2/5}$. As the discrepancy between the liquidity of asset increases (higher β), the estimator becomes less efficient.

Define for each $a = 1, \ldots, d$, $\mathcal{B}_{aa} = 0$ and

$$\mathcal{B}_{ab} = \frac{|k''(0)|\mathcal{A}^2}{2C_0^2}\int_0^{2\pi} e^{-itq}|\Sigma_{ab}(t)|dt \ ; \ \mathcal{A} = \lim_{n_1 \wedge n_2 \to \infty} \frac{n_1 \wedge n_2}{2\pi} \sup_{i,j} |t_i - s_j| 1_{\{I_{i,1} \cap I_{j,2} \neq \emptyset\}},$$

where $0 \leq \mathcal{A} < \infty$ under Assumption 2. $\mathcal{A}$ could be thought as a measure of the degree of non-synchronicity. When the two series are perfectly synchronized and balanced then $\mathcal{A} = 0$; otherwise it is $O(1)$ under Assumption 2.3. Define the asymptotic variance for the typical diagonal and off diagonal element:

$$\begin{aligned}
\mathcal{V}_{aa} &= 2C_0||k||^2 \int_0^{2\pi} e^{-i2qt}\Sigma_{aa}(t)d\mathcal{Q}_{aa}(t) \\
\mathcal{V}_{ab} &= C_0||k||^2 \int_0^{2\pi} e^{-i2qt} \left\{\Sigma_{aa}(t)\Sigma_{bb}(t)d\mathcal{Q}_{aabb}(t) + \Sigma_{ab}^2(t)d\mathcal{Q}_{abab}(t)\right\}.
\end{aligned}$$

The covariation between the integrated covariance estimator of asset a and b with the estimator of c and d is given by

$$\mathcal{V}_{ab,cd} = C_0\|k\|^2 \int_0^{2\pi} e^{-i2tq} \left\{\Sigma_{ac}(t)\Sigma_{bd}(t)d\mathcal{Q}_{acbd}(t) + \Sigma_{ad}\Sigma_{bc}(t)d\mathcal{Q}_{adbc}(t)\right\}.$$

Let $\mathcal{B}$ and $\mathcal{V}$ be the bias and covariance matrix of the vech of unique elements of our estimator. Define $\mathcal{D}_n^*$ to be the matrix of convergence rates, such that

$$\mathcal{D}_n = \text{diag}\left\{\text{vech}(\mathcal{D}_n^*)\right\} \ ; \ \{\mathcal{D}_n^*\}_{a,a} = \sqrt{n_a} \ ; \ \{\mathcal{D}_n^*\}_{a,b} = (n_a \wedge n_b)^{\vartheta}; \ \vartheta = \frac{4-2\beta}{5}$$

for $1 \leq \beta < 2$, where the upper bound $\vartheta = 2/5$ is obtained when the sample sizes are of the same order.

Theorem 1 *Suppose that Assumptions 1 - 4 hold. Then,*

$$\mathcal{D}_n \text{vech}\left\{\widehat{\mathcal{F}}(\mathbf{\Sigma})(q) - \mathcal{F}(\mathbf{\Sigma})(q)\right\} \Longrightarrow N\left(\mathcal{B}, \mathcal{V}\right).$$

remark When data is synchronized and balanced we have $\mathcal{B}_{ab} = 0$ and the covariation estimator achieves the same rate of convergence as the variance estimator. Our result is comparable with Malliavin and Mancino (2009) whose results were under sub-optimal bandwidth.

remark on efficiency If our goal is to achieve the most efficient estimator, we can estimate the asymptotic bias term and subtract it from our estimator. In that case we can get $\sqrt{n}$ convergence rate at the cost of sacrificing the positive definiteness of the estimator.

We also may estimate each element of the covariance matrix in most efficient way and use the clipping method to achieve p.s.d i.e. we can project a $d \times d$ symmetric covariance matrix estimate which has singular value decomposition, $\mathbf{U}^T diag[\lambda_1, \cdots, \lambda_p]\mathbf{U}$ as $\mathbf{U}^T diag[\lambda_1^+, \cdots, \lambda_p^+]\mathbf{U}$ where $\lambda_j^+ = \max\{\lambda_j, 0\}$. Whether we should emphasize on the efficiency of an estimator or on the covariance estimator that is guaranteed to be positive definite depends on problem at hand and we leave this choice to the practitioner.

REMARK ON DIFFERENCE FROM REALIZED KERNEL We analyze the asymptotic bias of our estimator v.s. the time domain estimator. Define $u_{ij} = t_i \wedge s_j$ and $l_{ij} = t_{i-1} \vee s_{j-1}$. Conditionally on $1_{\{i,j|u_{ij}>l_{ij}\}}$, the asymptotic bias of (11) is given by

$$\begin{aligned}
&\sum_{i,j} \int_{l_{i,j}}^{u_{i,j}} \{P_1(t) - P_1(l_{i,j})\} dP_2(t) + \int_{l_{i,j}}^{u_{i,j}} \{P_2(t) - P_2(l_{i,j})\} dP_1(t) \\
&+ \sum_{i,j} \left(\int_{l_{i,j}}^{u_{i,j}} - \int_0^{2\pi} \right) \Sigma_{12}(t) dt \\
&- \sum_{i,j} \int_{l_{i,j}}^{u_{i,j}} dP_1(t) \int_{l_{i,j}}^{u_{i,j}} dP_2(s) \{1 - k_H(t_i - s_j)\} = (i) + (ii) + (iii),
\end{aligned}$$

where $(i) = O_p(\{n_1 \vee n_2\}^{-1/2})$ and $(ii) = O_p(\{n_1 \vee n_2\}^{-1})$ are due to the discretization error of the continuous time signal, which depends inversely on the number of union of time stamps for two assets. $(iii) = O_p(\{\frac{n_1 \vee n_2}{H(n_1 \wedge n_2)}\}^2)$ is due to smoothing, which can be controlled as it depends on the bandwidth. The leading order term is then (iii). See Appendix. Our estimator does not suffer from the Epps effect, that is, a negative bias arising due to aligning the non synchronous data. In fact, the realized covariance applied to the refresh time aligned data has a non vanishing component (ii), which is analyzed in Theorem 1 of Zhang (2010). The stochastic bias term of the Realized Kernel can be derived similarly, given N as a refresh time sample size, the asymptotic bias is given by $O_p(N^{-1/2}) + O_p(N^{-1}) + O_p(H^{-2})$, where bandwidth H is chosen for the Realized Kernel. In practice, with high frequency data $N = o(n_1 + n_2)$. Hence, using all the realized transaction stamps is a much finer approximation for the real line $[0, 2\pi]$ than the coarser refresh time. If we let $N = (n_1 \wedge n_2)^r$, then under the optimal bandwidth, our estimator converges faster at $(n_1 \wedge n_2)^{(4-2\beta)/5}$ than the Realized Kernel at $(n_1 \wedge n_2)^{2r/5}$, when $r < 2-\ \beta$ i.e. $N = o(n_1 \wedge n_2)$. In two dimensional case, this condition will hold when two assets are traded very asynchronously and it will most likely hold when we are estimating the large dimensional covariance matrix.

4.2 With Microstructure Noise

The empirical evidence from the volatility signature plot suggests that the observed price deviates from the semimartingale assumption. More precisely various studies document that the observed high frequency returns have infinite quadratic variation. To model this phenomena, we make the following assumption.

Assumption 5. *Let $X_j(t_{i,j})$ be the observed log price of the j-th asset which has two additive components. One is a discrete realization of a continuous signal $P_j(t_{i,j})$ that satisfies the semimartingale Assumption 1 and the other component is a noise process with respect to the realization of transaction time $U_j(t_{i,j})$*

$$X_j(t_{i,j}) = P_j(t_{i,j}) + U_j(t_{i,j}). \tag{12}$$

In univariate studies, it is usually assumed that $U_j(t_{i,j})$ is a stationary time series, which has been supported by empirical studies. There has not been a lot of empirical work studying the cross autocorrelation of the microstructure noise for the multiple asset case. We think it is realistic to assume the following for the microstructure noise that allows cross-sectional correlation in the measurement error process.

Assumption 6. *Let $U_j(.), j = 1, \dots, d$ be a n dimensional stationary process, independent of the efficient price process with $E(U_j(.)) = 0$ and covariance function defined by $EU_a(t_{i,a})U_b(t_{j,b}) = \gamma(|t_{i,a} - t_{j,b}|/\widetilde{\Delta t}_{ab})$ that satisfies for some $d \times d$ p.s.d. covariance matrix $\mathbf{\Gamma}$ with (a, b)-th element denoted by Γ_{ab},*

$$(n_a \wedge n_b)^{-1} \sum_{i=1}^{n_a-1} \sum_{j=1}^{n_b-1} \gamma(|t_{i,a} - t_{j,b}|/\widetilde{\Delta t}_{ab}) \to \Gamma_{ab}.$$

We also assume that $|\mathrm{E}(U_a(t_{i,a})U_b(t_{j,b}), U_c(t_{r,c})U_d(t_{l,d}))| \leq \rho(M)$, *where* $M := \sup_{\{u,v\},\{p,s\}}\{(t_{u,p} - t_{v,s})/\widetilde{\Delta t}_{ps}\} < \infty$ and $\sum_{\nu}^{\infty} \rho(\nu)(1+\epsilon)^{\nu} < \infty$ *for some* $\epsilon > 0$.

We will show that our estimator in (8) is a consistent estimator of the Fourier transform of the covariance matrix even under the presence of microstructure noise. We add one further assumption on the end points.

ASSUMPTION 7. *The two end points, $X_j(t_{0,j})$ and $X_j(t_{n,j})$ are respectively an average of m_0 number of distinct observations on the interval $[t_{-1,j}, t_{0,j})$ and $[t_{n,j}, t_{n+1,j})$.*

This assumption turns about to be crucial for our estimator to achieve consistency. The time domain estimator by Barndorff-Nielsen et al. (2011) also assumes this condition. From Proposition 1 and 2 in the appendix, we derive the rate of convergence of our estimator by balancing the asymptotic variance of order $O_p(H/(n_1 \wedge n_2))$ and the asymptotic bias of order $O_p((n_1 \vee n_2)/H^2)$. The optimal bandwidth is given by $H = C_0 n^{\alpha^*}$, $\alpha^* = (2\beta+1)/5$, where $C_0 \in (0,\infty)$. When the two sample sizes are the same order i.e.($\beta = 1$), then $\alpha^* = 3/5$. In general, liquidity parameter $1 \le \beta < 2$ implies that $3/5 \le \alpha^* < 1$. The rate of convergence is then $(n_1 \wedge n_2)^{\vartheta}$, $0 < \vartheta := (2-\beta)/5 \le 1/5$, where the upper bound is achieved when the sample sizes are of the same order. We define a finite tuning parameter η in a following way. There exists $C^* \in (0,\infty)$ such that $n_\ell \sup_{0\le i\le n_\ell} \Delta t_{i,\ell} \le C^*$ for $\forall \ell = 1,\ldots,d$ under Assumption 2.1. We define $\eta = (\frac{C^*}{2\pi})^2$. Let denote $\mathcal{B} = \mathrm{vech}(C_0^{-2}\eta|k''(0)|\Gamma)$,and let $\mathcal{V}$ be as defined in Theorem 1. Let $\mathcal{D}_n^*$ be the matrix of convergence rates

$$\mathcal{D}_n = \mathrm{diag}\left\{\mathrm{vech}(\mathcal{D}_n^*)\right\} \quad \{\mathcal{D}_n^*\}_{a,b} = (n_a \wedge n_b)^{\vartheta} \quad \vartheta = \frac{2-\beta}{5}, 1 \le \beta < 2,$$

where the degree of liquidity parameter β is defined in Theorem 1. The the upper bound for ϑ is 1/5 which is obtained when $n_a/n_b = O(1)$.

THEOREM 2. *Suppose that Assumptions 1-7 hold. Then,*

$$\mathcal{D}_n \mathrm{vech}\left\{\widehat{\mathcal{F}}(\mathbf{\Sigma})(q) - \mathcal{F}(\mathbf{\Sigma})(q)\right\} \Longrightarrow N\left(\mathcal{B},\mathcal{V}\right).$$

The *instantaneous* covariance matrix is also a parameter of interest, see Kristensen (2010). We can construct an estimator of instantaneous covariation matrix by Fourier inverting the estimator given in (8)

$$\widehat{\Sigma}(t) = \frac{1}{2\pi} \sum_{|q| \le m/2} K_H(\lambda_q) \exp(iqt) \widehat{\mathcal{F}}(\mathbf{\Sigma})(q). \tag{13}$$

Suppose that the modulus of continuity of $\mathbf{\Sigma}(t)$ denoted by $\mathcal{C}(h)$ is given by

$$\mathcal{C}(h) := \sup_{|t-s| \le h} \|\mathbf{\Sigma}(t) - \mathbf{\Sigma}(s)\|_2. \tag{14}$$

The continuity assumption is met when each element of $\mathbf{\Sigma}(t)$ in Assumption 1 does not contain jumps, for example, when it is a Brownian semimartingale.

THEOREM 3. *Suppose that the assumptions of Theorem 2 hold and that* (14) *holds. Then, there exists a sequence* $\delta(n) \to 0$, *such that*

$$\lim_{n \to \infty} \sup_{\delta(n) \le t \le 2\pi - \delta(n)} \|\widehat{\mathbf{\Sigma}}(t) - \mathbf{\Sigma}(t)\|_2 = 0.$$

Often, practitioners encounter a problem of running a regression between variables that are asynchronously observed - for example we might be interested in the effect of returns and order book information of one asset on another asset. Hannan (1975) and Robinson (1975) are the earlier literature on using frequency domain to solve such problems. Mykland and Zhang (2006) discussed a general the set up of analysis of variance for continuous time regression.

5. NUMERICAL STUDY

5.1 Simulation studies

In the theoretical work, we assumed no leverage between the volatility and the return process. In the simulation studies, we relax this assumption and see if our estimator is robust to a presence of the leverage. The first DGP is same as the Brownian semimartingale with the perfect leverage given in Barndorff-Nielsen et al.(2011). For the second setting, we consider the stochastic volatility specified as a jump diffusion process given in Aït-Sahalia et al. (2010a) . We enrich their DGP by letting the instantaneous co-volatility coefficient to follow

Table 1: 2 dimensional covariation matrix simulation result ($\cdot/100$)

		Realized Cov			HY	Realized Kernel			**FRK**,$\overline{H}$			**FRK**,H^*		
NSR		(1,1)	(1,2)	(2,2)	(1,2)	(1,1)	(1,2)	(2,2)	(1,1)	(1,2)	(2,2)	(1,1)	(1,2)	(2,2)
DGP1 : Continuous Stochastic Volatility														
Sampling: (3/2,30) Balanced														
0	bias	1.0	(1.9)	0.8	1.0	1.0	(1.3)	**0.8**	0.9	**(0.6)**	1.4	0.9	**(0.6)**	1.4
	rmse	**8.0**	7.6	8.1	7.6	8.2	**7.5**	**8.0**	20.9	19.6	20.8	20.9	19.6	20.8
0.001	bias	23.2	7.3	22.9	10.3	1.6	0.9	1.6	1.0	0.8	1.7	**0.9**	**(0.1)**	**0.1**
	rmse	24.7	**10.6**	24.6	12.9	**16.7**	15.2	**15.4**	30.2	28.9	32.3	38.2	35.5	41.2
0.01	bias	225	92.3	225	96.4	3.5	2.3	4.0	**0.6**	**(1.2)**	**(1.5)**	**0.6**	**(1.2)**	**(1.5)**
	rmse	226	93.6	226	97.6	25.3	23.7	25.5	43.8	40.1	46.7	43.8	40.1	46.7
Sampling: (3/2,30) Unbalanced														
0	bias	0.1	(2.6)	**0.5**	0.4	**0.0**	(1.9)	0.6	1.0	(3.1)	1.4	1.1	**(0.1)**	1.3
	rmse	**6.7**	6.9	**7.0**	**6.6**	7.0	7.0	7.3	15.9	14.8	15.9	22.2	21.0	23.1
0.001	bias	23.5	6.6	21.3	9.5	2.0	**1.8**	3.1	1.5	(12.5)	5.0	**1.0**	(4.2)	**2.9**
	rmse	24.7	**9.6**	22.6	11.7	15.7	14.6	15.6	**10.5**	15.5	**12.2**	15.2	14.4	15.3
0.01	bias	239	92.5	210	95.7	2.6	2.3	**4.6**	1.9	(1.9)	11.6	**1.6**	**0.3**	4.9
	rmse	240	93.9	211	96.8	24.0	23.3	25.9	**18.1**	**16.7**	**21.7**	23.9	22.9	25.8
Sampling: (20,30) Unbalanced														
0	bias	**0.1**	(27.1)	0.7	0.8	0.2	(23.7)	**0.7**	0.9	**(0.4)**	1.8	0.9	**(0.4)**	1.8
	rmse	**8.3**	28.0	**8.2**	**9.3**	8.4	24.8	8.2	20.2	19.1	20.9	20.2	19.1	20.9
0.001	bias	16.8	(23.1)	16.2	4.9	2.0	0.1	2.8	1.1	**0.1**	1.9	**0.8**	0.4	**1.5**
	rmse	19.0	24.4	18.6	**11.0**	**16.3**	15.0	**16.2**	23.3	22.1	24.2	26.9	25.9	28.7
0.01	bias	162	10.4	155	41.1	3.8	2.2	4.7	1.7	0.6	1.9	**1.0**	**(0.5)**	**0.2**
	rmse	164	**16.5**	156	43.6	**24.3**	23.4	26.7	32.2	30.8	34.5	35.8	34.1	39.4
DGP2 : Jump Diffusion Stochastic Volatility														
Sampling: (3/2,30) Balanced														
0	bias	(0.3)	(2.0)	**(18.7)**	**(0.3)**	(0.2)	(1.5)	(18.6)	**(0.1)**	(0.7)	(18.8)	**(0.1)**	(0.7)	(18.8)
	rmse	**6.7**	**18.1**	20.3	18.6	7.1	18.3	**20.3**	18.0	23.4	24.6	18.0	23.4	24.6
0.001	bias	23.3	8.2	**4.9**	9.8	1.5	1.3	(16.9)	**0.0**	**0.1**	(19.3)	**0.0**	**0.1**	(19.3)
	rmse	24.7	**19.8**	**9.9**	21.1	**13.4**	22.0	21.8	29.1	29.2	31.9	29.1	29.2	31.9
0.01	bias	233	96.7	211	98.8	2.6	2.1	(15.8)	**1.4**	1.3	(18.0)	10.2	**(0.7)**	**(10.0)**
	rmse	234	99.6	212	102	19.8	24.5	25.3	42.6	38.9	42.8	**17.2**	**21.2**	**17.1**
Sampling: (3/2,30) Unbalanced														
0	bias	(0.3)	(1.9)	**(18.8)**	**(0.1)**	(0.2)	(1.6)	(18.8)	0.2	(2.5)	(19.1)	**0.0**	(0.5)	(18.9)
	rmse	7.1	**17.7**	**20.4**	18.2	**7.1**	17.9	20.5	13.7	21.1	23.0	20.1	24.4	25.3
0.001	bias	24.0	8.0	**2.6**	10.0	0.7	**0.5**	(17.5)	1.5	(8.3)	(14.6)	**0.7**	(3.2)	(17.5)
	rmse	25.5	19.7	**9.3**	21.1	14.0	20.7	21.5	9.2	**19.1**	18.0	**12.8**	20.5	21.6
0.01	bias	243	94.1	192	97.8	**1.4**	**1.0**	(15.7)	1.8	(1.6)	(8.6)	3.8	(3.3)	**2.6**
	rmse	244	97.1	193	101	22.1	24.7	23.8	15.2	21.6	17.3	**12.6**	**19.6**	**14.8**
Sampling: (20,30) Unbalanced														
0	bias	(0.1)	(16.7)	**(18.0)**	**(0.0)**	0.1	(14.7)	(18.0)	**(0.0)**	(0.9)	(18.9)	(0.2)	(0.1)	(18.9)
	rmse	**6.8**	21.5	**19.5**	**19.2**	6.8	20.4	19.6	17.6	23.3	24.2	24.2	27.0	28.3
0.001	bias	15.3	(13.2)	**(2.7)**	3.7	1.6	(0.4)	(17.9)	0.1	(0.4)	(18.5)	**0.1**	**(0.0)**	(18.8)
	rmse	17.1	**19.1**	**8.9**	19.6	**14.5**	21.2	21.7	19.7	24.5	25.1	23.9	26.7	28.0
0.01	bias	159	20.6	136	40.1	3.0	1.2	(15.9)	0.8	0.9	(17.8)	6.2	**0.7**	**(10.5)**
	rmse	**160**	26.8	137	45.5	22.7	25.8	24.1	30.2	30.6	32.6	17.9	**22.6**	**17.7**

the CIR process given in Barndorff-Nielsen and Shephard (2004). Specifically, for $j = 1, 2$-th asset, we let log price as $dP_j(t) = \sigma_j(t)dW_j(t)$ where $EdW_1(t)dW_2(t) = \rho_t dt$ such that $\rho(t) = (e^{2x(t)} - 1)/(e^{2x(t)} + 1)$ and $dx(t) = 0.03(0.64 - x(t))dt + 0.118x(t)dB_{xt}$ where B_{xt} is an independent Brownian motion. The DGP of Microstructure noise is a cross correlated AR(1) process formed with respect to a transaction time. This can be implemented by; $U_j(t_{i,j}) = \bar{U}_j(t_{i,j}) + \varepsilon(t_{i,j})$ with $\bar{U}_j(t_{i,j}) = \alpha\bar{U}_j(t_{i-1,j}) + \epsilon_j(t_{i,j})$, where idiosyncratic errors $\epsilon_j(t_{i,j})$ are independent Gaussian. The common disturbance is simulated by $\varepsilon_l \sim$ AR(1) for $\{T_l\}_{1\leq l\leq N_T} = \{t_{i,1} \cup t_{k,2}\}$. The variance of the noise is set to be proportionate to the sample integrated quarticity; $\zeta^2\sqrt{n_j{}^{-1}\sum_{i=1}^{n_j}\sigma_j^4(t_{i,j})}$, where $\zeta = \{0, 0.0.001, 0.01\}$ is a noise to signal ratio(NSR). We simulated the one second data assuming 6.5 hour daily trading, which give us 23,400 daily data points over 100 monte carlo sample. We designed the simulation to assess the impact of the asynchronicity on the estimator. We poisson sampled the data at the rate $\{(3/2, 30), (20, 30)\}$. The sampling rate $(3/2, 2)$ means that we sample the first asset on average per 1.5 second and the second asset per 2 second. To create a balanced sample for the rate $(3/2, 30)$, for the first asset, we sample on average at 1.5 second for the first half of the sample and at 30 second for the last half of the sample. For the second asset, we do this in reverse order. Then we have two assets that have the same number of transactions each day but traded very asynchronously.

For third simulation setting, we increase the dimension and consider a simple setting where log prices are given by $\mathbf{P}(t) = \mathbf{AB}(t)$ where $\mathbf{P}(t)$ is 10×1 vector of prices, $\mathbf{B}(t)$ is 3×1 independent Brownian motion and A is a factor loading matrix. This is poisson sampled at rate $\{2, 2, 4, 4, 8, 8, 10, 10, 30, 30\}$ and masked by the i.i.d gaussian noise.

Table 1 shows that the proposed estimator has the best bias profile and overall estimates the off-diagonal elements better than the other methods. We calculate the optimal bandwidth as given in Theorem 1 and 2 for each element of covariance matrix and take the minimum, maximum of these and average of the two. We report the results for average bandwidth,$\overline{H}$ and the bias minimizing bandwidth H^*. With carefully chosen bandwidth we can achieve

Table 2: Scalar function of 10 dimensional covariation matrix

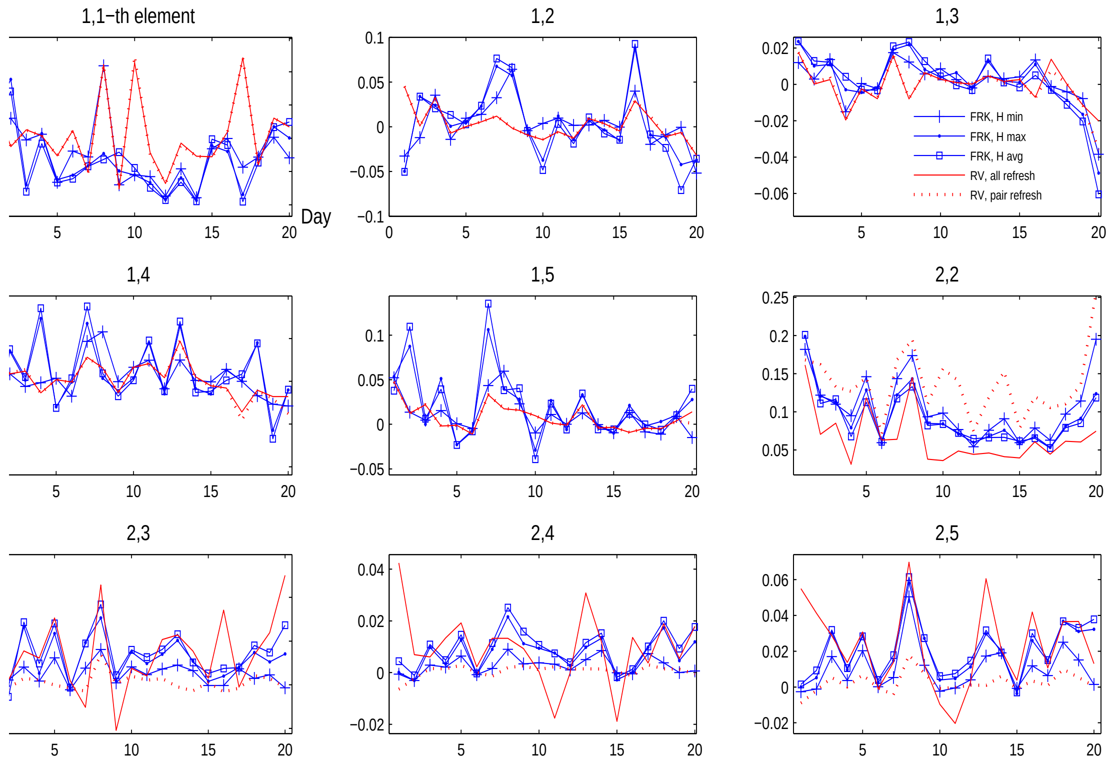
max(eigenvalue)		NSR=0		NSR=0.001		NSR=0.1	
		Bias	rMSE	Bias	rMSE	Bias	rMSE
RV_refresh		(2.34)	2.75	7.22	7.50	127	127
RV_fixed		**(0.85)**	3.18	0.40	3.28	15.4	16.5
Realized Kernel		(2.21)	2.65	0.38	3.00	1.29	4.82
Fourier RK	minH	(1.18)	**2.17**	**(0.28)**	**1.95**	1.22	**3.27**
	avgH			(0.47)	2.64	0.13	4.02
	maxH			(0.52)	3.13	**(0.03)**	4.92
portfolio							
RV_refresh		1.76	2.74	27.2	27.5	256	257
RV_fixed		**0.14**	4.09	4.14	6.31	40.4	42.0
Realized Kernel		1.66	2.67	1.67	4.16	3.23	6.86
Fourier RK	minH	0.26	**2.51**	3.88	4.87	6.88	8.55
	avgH			0.73	**3.67**	1.76	**5.58**
	maxH			**(0.20)**	3.99	**0.67**	5.96

the best root MSE under the presence of noise. When no noise is present, the Hayashi and Yoshida estimator performs well. The refresh time aligned method often performs better in estimating the integrated variance of the less traded asset; (2,2) element. This is since it effectively aligns data on the time stamps of the less traded asset. As analysed in the previous section, when there is no noise and the number of refresh time sample is size small, thr Realized Kernel underperforms in terms of bias. We observe also that the realized covariance estimator aligned on sparsely sampled data often performs well - this is because there is two opposing effect in terms of bias: the negative bias from epps effect and the positive bias from microstructure noise. The advantage of our estimator is most clear in estimating higher dimension covariance matrix as shown in Table 2. We estimate 10 dimensional integrated covariance matrix and compare the maximum of eigenvalues and variance of the equally weighted portfolio. We note that our estimator seems to have large variance, however it outperforms other methods under presence of microstructure noise.

5.2 Empirical Application

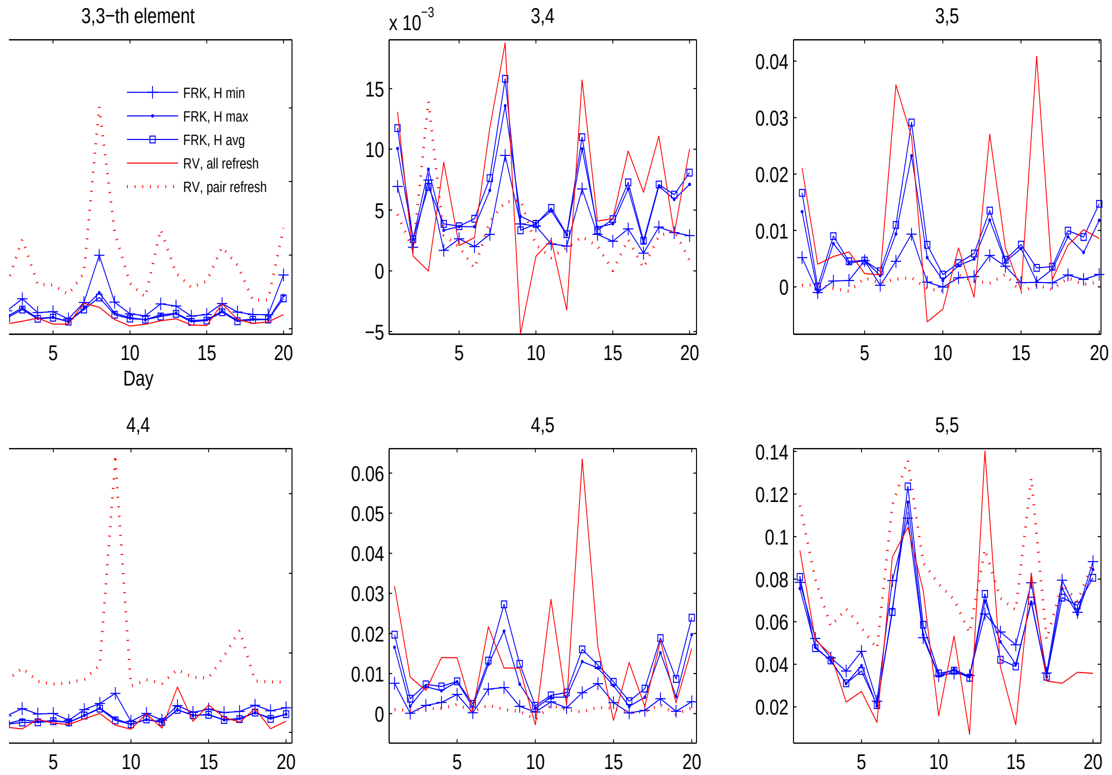
In this section we apply the Fourier Realized Kernel to a high frequency data. We analyzed five stocks in order of liquidity : Microsoft, J P Morgan, Dell, Caterpillar Inc. and Banco de Chile during 05-30 March 2007, where data is taken from WRDS TAQ database. We calculate the optimal bandwidth for individual asset by equalizing the squared bias and

Figure 2: Comparison between Realized Covariance and Fourier Realized Kernel : Estimates of $(1,1)$ to $(2,5)$ element of 5×5 matrix



the variance term given in Theorem 2. Let $n_\ell = n^{\beta_\ell}$ where n is a minimum of all sample sizes, then it is given by $H_\ell = \{\frac{\eta|k''(0)|}{\|k\|}\}^{2/5}\zeta_\ell^{4/5}n^{\frac{1+2\beta_\ell}{5}}$, where ζ_ℓ^2 is a squared noise to signal ratio for each asset. We applied maximum, minimum and average of the these individual bandwidths. The Figure 2 and Figure 3 compare the Realized Covariance and the proposed method in estimating the daily covariation matrix. Since the first asset is least traded, the all refresh time is effectively aligned on the trading time of the first asset. In estimating the integrated variance, the proposed method lies between the RV using pairwise refresh time (which will be dominated by the microstructure noise) and the RV using all refresh time (which is more sparsely sampled, therefore less affected by the noise). Most interesting case is the performance in estimating covariation for assets of different liquidity i.e. $(1,4)$ and

Figure 3: Comparison between Realized Covariance and Fourier Realized Kernel : Estimates of $(3,3)$ to $(5,5)$ element of 5×5 matrix



$(1,5)$-th element of the estimator in our case. The daily Realized Covariance take values closer to zero due to Epps effect whereas proposed estimator clearly gives us non trivial estimates.

APPENDIX A. PROOFS OF THEOREMS

We will prove the theorems for the general version of our estimator given in (8). We derive the results conditionally on the volatility matrix and the discretization time points therefore we regard these variables deterministic in the proofs. Throughout the proof we denote $C, C_1, C_2, \cdots$ finite constants.

Lemma 1. *Let* $\mathbf{P}(t)$ *defined on the filtered probability space* $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, \mathbb{P})$ *satisfies Assumption 1 with* $\boldsymbol{\mu}(t) = 0$. *Let* $f(t,s;q)$ be *a bounded and measurable function. Define* *square bracket operation to denote a quadratic covariation process defined in* (1). *Then,*

$$E\Big[\int_0^{2\pi}\int_0^{2\pi} f(t,s;q)dP_a(s)dP_b(t), \int_0^{2\pi}\int_0^{2\pi} f(t,s;q')dP_c(s)dP_d(t)\Big] \tag{A.1}$$
$$= \int_0^{2\pi}\int_0^{2\pi} \{f(t,s;q)f(t,s;q')d[P_a,P_c](s)d[P_b,P_d](t) + f(t,s;q)f(s,t;q')d[P_a,P_d](s)d[P_b,P_c](t)\}\,.$$

where double stochastic integral is Wiener-Itô sense.

Proof. The double Wiener-Itô integral can be written as

$$\int_0^{2\pi}\int_0^{2\pi} f(t,s;q)dP_a(s)dP_b(t) = \int_0^{2\pi}\int_0^{t} \{f(t,s;q)dP_a(s)dP_b(t) + f(s,t;q)dP_b(s)dP_a(t)\}\,,$$

so that the integrand is measurable with respect to $\mathcal{F}_t$ and the stochastic integration is well defined. Two terms above are martingale. Therefore (A.1) can be expressed as

$$\begin{bmatrix} \int_0^{2\pi}\int_0^t f(t,s;q)dP_a(s)dP_b(t) + \int_0^{2\pi}\int_0^t f(s,t;q)dP_b(s)dP_a(t), \\ \int_0^{2\pi}\int_0^t f(t,s;q')dP_c(s)dP_d(t) + \int_0^{2\pi}\int_0^t f(s,t;q')dP_d(s)dP_c(t) \end{bmatrix}.$$

Consider one of the cross product terms among four possible terms from above. By Itô's isometry,

$$E\Big[\int_0^{2\pi}\int_0^{t} f(t,s;q)dP_a(s)dP_b(t), \int_0^{2\pi}\int_0^{t} f(s,t;q')dP_d(s)dP_c(t)\Big] \tag{A.2}$$
$$= E\int_0^{2\pi}\left(\int_0^t f(t,s;q)dP_a(s)\right)\left(\int_0^t f(s,t;q')dP_d(s)\right)d[P_b,P_c](t).$$

where $d[P_b,P_c](t)$ means $[P_b,P_c]'(t)dt$, where the prime denotes the time derivative. By Fubini's theorem,

$$\int_0^{2\pi} E\left(\int_0^t f(t,s;q)dP_a(s)\right)\left(\int_0^t f(s,t;q')dP_d(s)\right)d[P_b,P_c](t).$$
$$= \int_0^{2\pi}\int_0^t f(t,s;q)f(s,t;q')d[P_a,P_d](s)d[P_b,P_c](t).$$

Together with the expected quadratic covariation of following terms,

$$E\Big[\int_0^{2\pi}\int_0^{t} f(s,t;q)dP_b(s)dP_a(t), \int_0^{2\pi}\int_0^{t} f(t,s;q')dP_c(s)dP_d(t)\Big]$$
$$= \int_0^{2\pi}\int_0^t f(s,t;q)f(t,s;q')d[P_b,P_c](s)d[P_a,P_d](t),$$

we have the result.

Lemma 2. *Define a step function*

$$\begin{aligned} f_n(t,s;q) &= \sum_{i,j} e^{-is_jq}e^{-i(t_i-s_j)\alpha}1_{[t_{i-1},t_i[}(t)1_{[s_{j-1},s_j[}(s)1_{\{I_{i,1}\cap I_{j,2}=\emptyset\}}(t,s) \\ g_n(t,s;q) &= \sum_{i,j} e^{-is_jq}e^{-i(t_i-s_j)\alpha}1_{[t_{i-1},t_i[}(t)1_{[s_{j-1},s_j[}(s). \end{aligned}$$

where discretization points $\{t_i, s_j\}$ satisfy Assumption 2.1. Then

$$\int_0^{2\pi}\int_0^{2\pi} f_n(t,s;q)dsdt = \int_0^{2\pi}\int_0^{2\pi} g_n(t,s;q)dsdt + O(\frac{1}{n_1\wedge n_2}).$$

Proof. Under Assumption 2.1,

$$\begin{aligned} &\int_0^{2\pi}\int_0^{2\pi}\sum_{i,j}\{1-1_{\{I_{i,1}\cap I_{j,2}=\emptyset\}}(t,s)\}e^{-it_jq}e^{-i(t_i-s_j)\alpha}1_{[t_{i-1},t_i[}(t)1_{[s_{j-1},s_j[}(s)dsdt \\ &= \int_0^{2\pi}\int_0^{2\pi}\sum_{i,j} e^{-is_jq}e^{-i(t_i-s_j)\alpha}1_{[t_{i-1},t_i[}(t)1_{[s_{j-1},s_j[}(s)1_{\{I_{i,1}\cap I_{j,2}\neq\emptyset\}}(t,s)dsdt \\ &\le \sup_i(\Delta t_i)\sup_j(\Delta s_j)\sum_{i,j}1_{\{I_{i,1}\cap I_{j,2}\neq\emptyset\}} = O(\frac{1}{n_1\wedge n_2}). \end{aligned}$$

Lemma 3. *Define a step function weighted by kernel satisfying Assumption 3 by*

$$f_n(t,s;q,a,b) = \sum_{i,j} e^{-it_{j,b}q}k_H(t_{i,a}-t_{j,b})1_{[t_{i-1,a},t_{i,a}[}(t)1_{[t_{j-1,b},t_{j,b}[}(s)1_{\{I_{i,a}\cap I_{j,b}=\emptyset\}}(t,s), \quad \text{(A.3)}$$

$$g_n(t,s;q,a,b) = \sum_{i,j} e^{-it_{j,b}q}k_H(t_{i,a}-t_{j,b})1_{[t_{i-1,a},t_{i,a}[}(t)1_{[t_{j-1,b},t_{j,b}[}(s).$$

Then it holds that

$$\begin{aligned} &\frac{n_a\wedge n_b\wedge n_c\wedge n_d}{H}\int_0^{2\pi}\int_0^{2\pi} f_n(t,s;q,a,b)f_n(t,s;q,c,d)d[P_a,P_c](s)d[P_b,P_d](t) \quad \text{(A.4)} \\ &\to \|k\|^2\int_0^{2\pi} e^{-i2tq}[P_a,P_c]'(t)[P_b,P_d]'(t)d\mathcal{Q}_{acbd}(t), \end{aligned}$$

where $\mathcal{Q}_{\cdot}(t)$ is defined in Assumption 2.

Proof. We will proceed by proving for univariate, bivariate and general version of the formula in (A.4). First note that for any function $d(\cdot,\cdot)$, it holds that $\sum_{i,j=1}^n d(i,j) = \sum_{h=0}^{n-1}\sum_{j=1}^{n-h} d(j,j+h) + \sum_{h=1}^{n-1}\sum_{j=1+h}^{n} d(j,j-h)$, which we will denote by $\sum_{i,j=1}^n d(i,j) =$

$\sum_{h=0}^{n-1}\sum_{j=1}^{n-h} d(j,j+h)[2]$. For univariate case, by Lemma 2, we can replace f_n by g_n with error $O(n^{-1})$. Then (A.4) is approximated by $\frac{n}{H}\sum_{h=0}^{n-1}\sum_{j=1}^{n-h}(e^{-it_j2q}+e^{-it_jq}e^{-it_{j+h}q})k_H^2(t_{j+h}-t_j)[P,P]'(t_{j+h})[P,P]'(t_j)\Delta t_{j+h}\Delta t_j[2]$, which is equal to

$$\frac{n}{H}\sum_{h=0}^{n-1}k_H^2(t_h)\sum_{j=1}^{n-h}(e^{-it_j2q}+e^{-it_jq}e^{-it_{j+h}q})[P,P]'(t_{j+h})[P,P]'(t_j)\Delta t_{j+h}\Delta t_j[2] \tag{A.5}$$

In (A.5) we can make approximation $t_i - t_{i-h} \simeq t_h$ under Assumption 2. We may replace the second summation in (A.5) to run to n, with approximation error

$$\begin{aligned}
&\frac{n}{H}\sum_{h=0}^{n-1}k_H^2(t_h)\sum_{j=1}^{h}(e^{-it_j2q}+e^{-it_jq}e^{-it_{j+h}q})[P,P]'(t_{j+h})[P,P]'(t_j)\Delta t_{j+h}\Delta t_j[2]\\
\le\;& C_2(\sup\Delta t_j)^2 n\sum_{|h|<n}\frac{h}{H}k_H^2(t_h)\simeq\frac{H}{n}C_2\int_{-\infty}^{\infty}xk^2(x)dx.
\end{aligned}$$

Then (A.5) is equal to

$$\begin{aligned}
&\frac{n}{H}\left\{k_H^2(0)+2\sum_{h=1}^{n-1}k_H^2(t_h)\right\}\sum_{j=1}^{n}(e^{-it_j2q}+e^{-it_jq}e^{-it_{j+h}q})[P,P]'(t_{j+h})[P,P]'(t_j)\\
&\times\Delta t_{j+h}\Delta t_j+O(\frac{H}{n})\to 2\|k\|^2\int_0^{2\pi}e^{-it2q}([P,P]')^2(t)d\mathcal{Q}_{11}(t),
\end{aligned}\tag{A.6}$$

where $\mathcal{Q}_{11}(t)$ is defined in Assumption 2. For the bivariate case we first establish some inequalities regarding the two grids of time stamps,

$$\sum_{i,j}\Delta t_i\Delta s_j\ge\sum_{i,j}\Delta t_i\Delta s_jk_H^2(t_i-s_j)\ge\sum_{i,j}\Delta t_i\Delta s_j1_{\{I_{i,1}\cap I_{j,2}\neq\emptyset\}}, \tag{A.7}$$

since they are of order $O(1)$, $O(H/(n_1\wedge n_2))$ and $O((n_1\wedge n_2)^{-1})$ respectively. Recalling that $\{T_l\}_{1\le l\le N_T}$ are union of time stamps, the lower bound for all (A.7) is $\sum_{i,j}(t_i\wedge s_j - t_i\vee s_j)^2\le\sup_l|T_l-T_{l-1}|\sum_{1\le l\le N_T}|T_l-T_{l-1}| = O((n_1\vee n_2)^{-1})$. By Riemann approximation of a continuous integral and (A.7), under Assumption 2.2,

$$\begin{aligned}
&(n_1\wedge n_2)\sum_{i,j}e^{-i2s_jq}[P_1,P_1]'(t_i)[P_2,P_2]'(s_j)\Delta t_i\Delta s_j1_{\{I_{i,1}\cap I_{j,2}\neq\emptyset\}}\\
\to&\int_0^{2\pi}e^{-i2tq}[P_1,P_1]'(t)[P_2,P_2]'(t)d\mathcal{Q}_{1122}(t).
\end{aligned}$$

Using Lemma 3, then it holds that

$$\frac{n_1 \wedge n_2}{H} \sum_{i,j} e^{-i2s_j q} k_H^2(t_i - s_j) \int_{t_{i-1}}^{t_i} d[P_1, P_1](t) \int_{s_{j-1}}^{s_j} d[P_2, P_2](s)$$

$$\simeq \frac{n_1 \wedge n_2}{H} \sum_{i,j} e^{-i2s_j q} k_H^2(t_i - s_j)[P_1, P_1]'(t_i)[P_2, P_2]'(s_j)\Delta t_i \Delta s_j$$

$$\simeq \|k\|^2 (n_1 \wedge n_2) \sum_{i,j} e^{-i2s_j q}[P_1, P_1]'(t_i)[P_2, P_2]'(s_j)\Delta t_i \Delta s_j 1_{\{I_{i,1} \cap I_{j,2} \neq \emptyset\}}$$

$$\to \|k\|^2 \int_0^{2\pi} e^{-i2tq}[P_1, P_1]'(t)[P_2, P_2]'(t) d\mathcal{Q}_{1122}(t), \tag{A.8}$$

where the approximation errors are similarly derived as the univariate case. Next we note that the cross product term simplifies to $g_n(t, s; q)g_n(s, t; q) = \sum_{i,j,k,l} e^{-is_j q} e^{-is_l q} k_H(t_i - s_j)k_H(t_k - s_l)1_{[t_{i-1} \vee s_{l-1}, t_i \wedge s_l[}(t)1_{[t_{k-1} \vee s_{j-1}, t_k \wedge s_j[}(s)$. Then by Lemma 3 and (A.7),

$$\frac{n_1 \wedge n_2}{H} \int_0^{2\pi} \int_0^{2\pi} f_n(t, s; q) f_n(s, t; q) d[P_1, P_2](t) d[P_2, P_1](s)$$

$$\simeq \frac{n_1 \wedge n_2}{H} \sum_{i,j,k,l} e^{-is_j q} e^{-is_l q} k_H(t_i - s_j) k_H(t_k - s_l)[P_1, P_2]'(t_i \wedge s_l)[P_2, P_1]'(t_k \wedge s_j)$$

$$\times (t_i \wedge s_l - t_{i-1} \vee s_{l-1})(t_k \wedge s_j - t_{k-1} \vee s_{j-1}) 1_{\{I_{i,1} \cap I_{l,2} \neq \emptyset\}} 1_{\{I_{k,1} \cap I_{j,2} \neq \emptyset\}}$$

$$\simeq \|k\|^2 (n_1 \wedge n_2) \sum_{i,j,l} e^{-is_j q} e^{-is_l q}[P_1, P_2]'(t_i \wedge s_l)[P_2, P_1]'(t_i \wedge s_j)$$

$$\times (t_i \wedge s_l - t_{i-1} \vee s_{l-1})(t_i \wedge s_j - t_{i-1} \vee s_{j-1}) 1_{\{I_{i,1} \cap I_{l,2} \neq \emptyset\}} 1_{\{I_{i,1} \cap I_{j,2} \neq \emptyset\}}$$

$$\to \|k\|^2 \int_0^{2\pi} e^{-i2tq}([P_1, P_2]')^2(t) d\mathcal{Q}_{1212}(t). \tag{A.9}$$

To analyse (A.4) for four assets, we first replace off-diagonal step function f_n by g_n by Lemma 2 and note that $\int_0^{2\pi} \int_0^{2\pi} g_n(t, s; q, a, b) g_n(t, s; q, c, d) ds dt$ is given by

$$\sum_{i,j,k,l} k_H(t_{i,a} - t_{j,b}) k_H(t_{k,c} - t_{l,d})(t_{i,a} \wedge t_{k,c} - t_{i-1,a} \vee t_{k-1,c})(t_{j,b} \wedge t_{l,d} - t_{j-1,b} \vee t_{l-1,d}) \tag{A.10}$$

The upper bound for (A.10) is given by

$$\sum_{i,j,k,l} (t_{i,a} \wedge t_{k,c} - t_{i-1,a} \vee t_{k-1,c})(t_{j,b} \wedge t_{l,d} - t_{j-1,b} \vee t_{l-1,d}) 1_{\{I_{i,a} \cap I_{k,c} \neq \emptyset\}} 1_{\{I_{j,b} \cap I_{l,d} \neq \emptyset\}} = O(1). \tag{A.11}$$

The lower bound for (A.10) is given by

$$\sum_{i,j,k,l}(t_{i,a}\wedge t_{k,c}-t_{i-1,a}\vee t_{k-1,c})(t_{j,b}\wedge t_{l,d}-t_{j-1,b}\vee t_{l-1,d})1_{\{I_{i,a}\cap I_{j,b}\cap I_{k,c}\cap I_{l,d}\neq\emptyset\}} \tag{A.12}$$
$$\leq C\frac{1}{n_a\vee n_c}\frac{1}{n_b\vee n_d}\sum_{i,j,k,l}1_{\{I_{i,a}\cap I_{j,b}\cap I_{k,c}\cap I_{l,d}\neq\emptyset\}}=O(\frac{n_a\vee n_b\vee n_c\vee n_d}{(n_a\vee n_c)(n_b\vee n_d)}),$$

which will be order of inverse of second or third largest sample size. It is bigger or equal to sum of squared union of all four time stamps which has order $O((n_a\vee n_b\vee n_c\vee n_d)^{-1})$. Then we may construct a quantity that lies between the (A.12) and (A.11) as following way. For simplicity assume that $n_a<n_b<n_c<n_d$, then

$$\sum_{i,j,k,l:t_{i,a},t_{j,b},t_{k,c},t_{l,d}<t}(t_{i,a}\wedge t_{j,b}-t_{i-1,a}\vee t_{j-1,b})(t_{k,c}\wedge t_{l,d}-t_{k-1,c}\vee t_{l-1,d})$$
$$\times 1_{\{I_{i,a}\cap I_{j,b}\neq\emptyset\}}1_{\{I_{k,c}\cap I_{l,d}\neq\emptyset\}}1_{\{I_{min\{(i,a),(j,b)\}}\cap I_{min\{(k,c),(l,d)\}}\neq\emptyset\}}$$
$$\leq C\frac{1}{n_b}\frac{1}{n_d}\sum_{i=1}^{n_a}\sum_{k=1}^{n_c}\sharp\{t_{k,c}\in[t_{i-1,a},t_{i,a}[\}\sum_{j=1}^{n_b}\sharp\{t_{j,b}\in[t_{i-1,a},t_{i,a}[\}\sum_{l=1}^{n_d}\sharp\{t_{l,d}\in[t_{k-1,c},t_{k,c}[\},$$

which is order of $C\frac{1}{n_b}\frac{1}{n_d}n_a\frac{n_c}{n_a}\frac{n_b}{n_a}\frac{n_d}{n_c}=O({n_a}^{-1})$ under Assumption 2.3. Then (A.4) is derived similarly as (A.9) using Lemma 3 and inequalities involving the quadratic variation of time derived above.

Lemma 4. *Let $\mathbf{P}(t)$ defined on probability space $(\Omega,\mathcal{F},\mathcal{F}_{t\geq 0},\mathbb{P})$ satisfying the Assumption 1 and let sub-$\boldsymbol{\sigma}$ field of $\mathcal{F}$ by $\mathcal{G}=\boldsymbol{\sigma}(\mathbf{P})$. The Z is a standard normal variable on the suitable extension of probability space and $\mathcal{V}$ is a $\mathcal{G}$-measurable stochastic variance. Then it holds that for $f_n(\cdot)$ given in (A.3),*

$$\sqrt{\frac{n}{H}}\int_0^{2\pi}\int_0^{2\pi}f_n(t,s;q)dP_1(s)dP_2(t)\Longrightarrow\sqrt{\mathcal{V}}Z.$$

where convergence is $\mathcal{G}$-stably in law.

Proof. Stable convergence is notion of joint convergence and stronger than convergence in law. See Aldous and Eagleson (1978) Proposition 1 for the definition of a stable convergence. Let the discretized filtration by $\mathcal{F}_i$, $i=\max_j\{t_j\leq t\}$. For the discretized sequence $\chi_i^n=\sqrt{\frac{n}{H}}\Delta P_1(t_i)\sum_{j:s_j<t_i}\Delta P_2(s_j)k_H(t_i-s_j)e^{-is_jq}$ which is adopted to $\mathcal{F}_i$, we

show the stable convergence of $Z_t^n := \sum_{\max_i\{t_i \le t\}} \chi_i^n$ to $Z_t = \int_0^t v_s dW_s$, a $\mathcal{F}_t$-conditional Gaussian martingale. Under the following conditions: (1) $\sum_i E(|\chi_i^n|^2|\mathcal{F}_{i-1}) \to_p [Z,Z]_t$; (2) $\sum_i^n E(|\chi_i|^2 1_{\{|\chi_i^n|>\epsilon\}}|\mathcal{F}_{i-1}) \to_p 0 \ \forall\epsilon$; we have $Z^n \Longrightarrow Z$ stably. See the proof for Theorem 3.2 in Jacod (1997). The sufficient condition for the conditional Lindberg condition in (2) is the Liapanov condition $\sum_i E(\{\chi_i^n\}^{2+\varepsilon}|\mathcal{F}_{i-1}) \to_p 0$,for $\varepsilon > 0$. We will show for $\varepsilon = 2$ in the proofs for Theorem 1 and Theorem 2.

A.1 Proof of Theorem 1

We first prove for the diagonal element. Consider the first element of the centered estimator

$$\mathcal{E} = \sum_{|\alpha|\le m/2} K_H(\lambda_\alpha)\mathcal{F}_n(P_1)(\alpha)\mathcal{F}_n(P_1)(q-\alpha) - \mathcal{F}(\Sigma_{11})(q).$$

We drop the subscript denoting asset for now. We can decompose the centered estimator into two terms, $\mathcal{E} = M_1 + M_2$:

$$M_1 = \sum_{i=1}^n \Delta P^2(t_i)e^{-it_iq} - \int_0^{2\pi} e^{-iqt}d[P,P](t) \quad ; M_2 = \sum_{i\neq j} \Delta P(t_i)\Delta P(t_j)k_H(t_i-t_j)e^{-it_jq}.$$

We will show that $\sqrt{\frac{n}{H}}M_1 = o_p(1)$ and $\sqrt{\frac{n}{H}}M_2$ stably converges to a zero mean Gaussian variable By Itô's formula, M_1 can be further decomposed into a martingale $M_{11} = 2\sum_{i=0}^n \int_{t_{i-1}}^{t_i} \{P(t)-P(t_{i-1})\}e^{-ikt_i}dP(t) = O_p(n^{-1/2})$ and a predictable finite variation component $A = \sum_{i=1}^n \int_{t_{i-1}}^{t_i}(e^{-itk} - e^{-it_ik})d[P,P](t) = O_p(n^{-1})$. This is the Euler discretization error and its distribution is given by the Theorem 5.5 of Jacod and Protter (1998). Therefore, $\sqrt{\frac{n}{H}}M_1 = O_p(H^{-1/2}) = o_p(1)$. Given the off-diagonal step function $f_n(t,s;q)$ in (A.3), we can express $M_2 = \int_0^{2\pi}\int_0^{2\pi} f_n(t,s;q)dP(s)dP(t)$ which has zero expectation. Then it holds by Lemma 1, $E[M_2,M_2] = 2E\int_0^{2\pi}\int_0^{2\pi} f_n^2(t,s;q)dP(s)dP(t)$, which is equal to (A.6) in Lemma 3. To verify the condition (2) in Lemma 4, let $\chi_i^n = \{\sum_{j<i}\sqrt{\frac{n}{H}}\Delta P(t_i)\Delta P(t_j)k_H(t_i - t_j)\,(e^{-it_jq} + e^{-it_iq})\}$. Then, $E|\chi_i^n|^4$ for $i = n$ is bounded by $2^4\times$

$$\begin{aligned} &n^2H^{-2}E\{\sum_{h=1}^n \Delta P(t_i)\Delta P(t_{i-h})k_H\,(t_h)\}^4 \\ &= 9n^2H^{-2}\sum_{h=1}^n E\left(\int_{t_{i-1}}^{t_i}[P,P]'(t)dt\right)^2\left(\int_{t_{i-h-1}}^{t_{i-h}}[P,P]'(t)dt\right)^2 k_H^4\,(t_h) + 18n^2H^{-2}\end{aligned}$$

$$\times \sum_{h,l=1}^{n} E\left(\int_{t_{i-1}}^{t_i} [P,P]'(t)dt\right)^2 \int_{t_{i-h-1}}^{t_{i-h}} [P,P]'(t)dt \int_{t_{i-l-1}}^{t_{i-l}} [P,P]'(t)dt k_H^2\left(t_h\right) k_H^2\left(t_l\right)$$

$$\leq 9n^2H^{-1} \sup_t \ ([P,P]'(t))^4 \sup_i(\Delta {t_i}^4) \times (\frac{1}{H}\sum_{h=1}^{n} k_H^4\left(t_h\right) + \frac{2}{H}\sum_{h,l=1}^{n} k_H^2\left(t_h\right) k_H^2(t_l)$$

$$= n^{-2}H^{-1}C_1\int_0^\infty k^4(x)dx + n^{-2}C_2(\int_0^\infty k^2(x)dx)^2 = O_p(n^{-2}).$$

where $[P,P]'(t)dt = d[P,P](t)$. In univariate case this simplifies to $[P,P]'(t) = \sigma^2(t)$. The pen-ultimate equality is using Assumption 2.1. Therefore the condition (2) in Lemma 4 is satisfied.

We now give a result for the off-diagonal element of the estimator when time stamps are asynchronous and sample sizes are unbalanced. We first show for the bivariate case and will extend the result to higher dimension. Denote the transaction time of the first asset $t_{i,1} = t_i$ and the second asset $t_{j,2} = s_j$. The centered estimator in (8) can be decomposed into, $\mathcal{E} = M_1 + M_2$, where

$$M_1 = \sum_{i,j} e^{-is_jq}k_H(t_i - s_j)\Delta P_1(t_i)\Delta P_2(s_j)1_{\{I_{i,1}\cap I_{j,2}\neq\emptyset\}} - \int_0^{2\pi} e^{-iqt}d[P_1,P_2](t)$$

$$M_2 = \sum_{i,j} e^{-is_jq}k_H(t_i - s_j)\Delta P_1(t_i)\Delta P_2(s_j)1_{\{I_{i,1}\cap I_{j,2}=\emptyset\}}.$$

We first derive the asymptotic bias. Let $u_{ij} = t_i \wedge s_j$ and $l_{ij} = t_{i-1} \vee s_{j-1}$. Then,

$$E(M_1) = E(\sum_{i,j} e^{-is_jq}\int_{l_{i,j}}^{u_{i,j}} dP_1(t)\int_{l_{i,j}}^{u_{i,j}} dP_2(s)1_{\{I_{i,1}\cap I_{j,2}\neq\emptyset\}} - \int_0^{2\pi} e^{-iqt}d[P_1,P_2](t))$$
$$- E(\sum_{i,j} e^{-is_jq}\int_{l_{i,j}}^{u_{i,j}} dP_1(t)\int_{l_{i,j}}^{u_{i,j}} dP_2(s)\{1 - k_H(t_i - s_j)\}1_{\{I_{i,1}\cap I_{j,2}\neq\emptyset\}}).$$

By multivariate Itô's calculus, $E(M_1)$ is given by the expectation of following terms conditionally on $1_{\{I_{i,1}\cap I_{j,2}\neq\emptyset\}}$

$$\sum_{i,j} e^{-is_jq}\int_{l_{i,j}}^{u_{i,j}} \{P_1(t) - P_1(l_{i,j})\}dP_2(t) + e^{-is_jq}\int_{l_{i,j}}^{u_{i,j}} \{P_2(t) - P_2(l_{i,j})\}dP_1(t) \qquad \text{(A.13)}$$

$$+ \sum_{i,j}\int_{l_{i,j}}^{u_{i,j}} (e^{-is_jq} - e^{-itq})d[P_1,P_2](t) \qquad \text{(A.14)}$$

$$- \sum_{i,j} e^{-is_jq}\int_{l_{i,j}}^{u_{i,j}} dP_1(t)\int_{l_{i,j}}^{u_{i,j}} dP_2(s)\{1 - k_H(t_i - s_j)\}. \qquad \text{(A.15)}$$

Recalling the definition of the union of time stamps in Assumption 2, the order of magnitude of the first term in (A.13) is given by

$$
\begin{aligned}
&\sum_{1\le i\le n_1, 1\le j\le n_2} e^{-is_jq}\int_{l_{i,j}}^{u_{i,j}}\{P_1(t)-P_1(l_{i,j})\}dP_2(t)\\
&=\sum_{l=1}^{N_T}\int_{T_{l-1}}^{T_l}\{P_1(t)-P_1(T_{l-1})\}dP_2(t)-\sum_{i,j}(1-e^{-is_jq})\int_{l_{i,j}}^{u_{i,j}}\{P_1(t)-P_1(l_{i,j})\}dP_2(t)\\
&=O_p(N_T^{-1/2})+O_p(n_2^{-1}N_T^{-1/2}).
\end{aligned}
$$

The order of the magnitude for the second term in (A.13) is derived in a similar way. The change of discretization points to the union of the time points are without error and holds analytically. The order of (A.14) is $O_p(n_2^{-1})$ since we can replace summation $\sum_{i,j}\int_{l_{i,j}}^{u_{i,j}}$ by $\sum_{1\le j\le n_2}\int_{s_{j-1}}^{s_j}$. This term is zero for an integrated (co)variance estimator, $q=0$. The asymptotic bias term conditional on the volatility path is given by

$$
\begin{aligned}
&E\sum_{i,j}e^{-is_jq}\int_{l_{i,j}}^{u_{i,j}}dP_1(t)\int_{l_{i,j}}^{u_{i,j}}dP_2(s)\{1-k_H(t_i-s_j)\}1_{\{I_{i,1}\cap I_{j,2}\neq\emptyset\}}\\
&\simeq\sum_{i,j}e^{-is_jq}\int_{l_{i,j}}^{u_{i,j}}d[P_1,P_2](t)\{-\frac{1}{2}k''(0)\left(\frac{(t_i-s_j)}{\widetilde{\Delta t}H}\right)^2\}1_{\{I_{i,1}\cap I_{j,2}\neq\emptyset\}}\\
&\le\left(\frac{n_1\vee n_2}{(n_1\wedge n_2)H}\right)^2\frac{1}{2}\left\{\frac{n_1\wedge n_2}{2\pi}\sup_{i,j}|t_i-s_j|1_{\{I_{i,1}\cap I_{j,2}\neq\emptyset\}}\right\}^2|k''(0)|\sum_{i,j}e^{-is_jq}\int_{l_{i,j}}^{u_{i,j}}d\left|[P_1,P_2]\right|(t)\\
&=\left(\frac{n_1\vee n_2}{(n_1\wedge n_2)H}\right)^2\frac{1}{2}\mathcal{A}^2|k''(0)|\int_0^{2\pi}e^{-itq}d\left|[P_1,P_2]\right|(t).
\end{aligned}
$$

by Taylor expansion of $\{k(0)-k_H(t_i-s_j)\}$, the first approximation holds by Assumption 2.3 and $k'(0)=0$. Then the order of the stochastic bias M_1 is given by $O_p(N_T^{-1/2})+O_p(n_2^{-1})+O_p(\{\frac{n_1\vee n_2}{(n_1\wedge n_2)H}\}^2)$ for estimator at non-zero frequency and $O_p(N_T^{-1/2})+O_p(\{\frac{n_1\vee n_2}{(n_1\wedge n_2)H}\}^2)$ for integrated (co)variance estimator. In both cases, the leading order term for the bias is the last term under the optimal bandwidth.

We next analyze M_2 which can be expressed as

$$
M_2=\int_0^{2\pi}\int_{s<t}f_n(t,s;q)dP_2(s)dP_1(t)+\int_0^{2\pi}\int_{s<t}f_n(s,t;q)dP_1(s)dP_2(t),
$$

where $f_n(t,s;q)$ is given in (A.3). It has a zero expectation and by Lemma 1 and Lemma 3, $E\left[M_2, M_2\right]$ multiplied by the appropriate rate of convergence is equal to (A.8)+(A.9). To complete the proof for the stable convergence, define $\chi_i^n = \sum_{j:s_j<t_i} \sqrt{\frac{n}{H}}\Delta P_1(t_i)\Delta P_2(s_j)k_H(t_i - s_j)e^{-is_jq}1_{\{I_{i,1}\cap I_{j,2}=\emptyset\}}$. Then $\sup_i E|\chi_i^n|^4 = O((n_1\wedge n_2)^{-2})$ which can be proved similarly as the univariate case. Therefore the condition (2) in Lemma 4 is met. To show a convergence of covariation matrix estimator to a multivariate Gaussian distribution by a Cramer-Wold device, it is sufficient and necessary to show that the linear combination of the elements of the matrix estimator converges to a univariate Gaussian random variable. Let denote $\mathcal{R}(q) := \widehat{\mathcal{F}}(\mathbf{\Sigma})(q) - \mathcal{F}(\mathbf{\Sigma})(q)$ and consider the linear combination of the element $\mathbf{a}^\top\mathcal{R}(q)\mathbf{b}$ and $\mathbf{c}^\top\mathcal{R}(q)\mathbf{d}$. Note that $\mathbf{a}^\top\mathcal{R}(q)\mathbf{c}\mathbf{b}^\top\mathcal{R}(q)\mathbf{d} =$vech$(\mathbf{a}\mathbf{b}^\top)^\top(\mathcal{R}(q)\otimes\mathcal{R}(q))$vech$(\mathbf{d}\mathbf{c}^\top)$. The expectation of the above expression depends on $E\{\mathcal{R}(q)\otimes\mathcal{R}(q)\}$. Each element of this is given in Lemma 3.

A.2 Proof of Theorem 2

We analyze our estimator applied to microstructure noise not affected by the end points U_0 and U_n, given by

$$\hbar^{-2}\sum_{i=1}^{n_1-1}\sum_{j=1}^{n_2-1} U_1(t_i)U_2(s_j)e^{-is_j\alpha}k_H^{''}(t_i - s_j)\Delta t_{i+1}\Delta s_{j+1} \tag{A.16}$$

$$+\,\hbar^{-1}\sum_{i=1}^{n_1-1}\sum_{j=1}^{n_2-1} U_1(t_i)U_2(s_j)e^{-is_j\alpha}(e^{-i\Delta s_{j+1}\alpha}-1)k_H^{'}(t_i - s_j)\Delta t_{i+1}. \tag{A.17}$$

The upper bound for expectations of (A.17) is given by.

$$\begin{aligned}
&\hbar^{-1}\sup_j\left|1-e^{-i\Delta s_{j+1}\alpha}\right|\sup_i(\Delta t_{i+1})\left|\sum_{i=1}^{n_1-1}\sum_{j=1}^{n_2-1}E\{U_1(t_i)U_2(s_j)\}k_H^{'}(t_i-s_j)\right|\\
&= C_1\frac{n_1\vee n_2}{H}\frac{1}{n_1n_2}\left|\{\sum_{|t_i-s_j|/\widetilde{\Delta t}\le\sqrt{H}} + \sum_{|t_i-s_j|/\widetilde{\Delta t}>\sqrt{H}}\}E\{U_1(t_i)U_2(s_j)\}k_H^{'}(t_i-s_j)\right|\\
&\le C_1(n_1\wedge n_2)^{-1}H^{-1}\{\sup_{|t_i-s_j|/\widetilde{\Delta t}\le\sqrt{H}}|k_H^{'}(t_i-s_j)|\left|\sum_{|t_i-s_j|/\widetilde{\Delta t}\le\sqrt{H}}\gamma(\{t_i-s_j\}/\widetilde{\Delta t})\right|
\end{aligned}$$

$$+ \sup_{|t_i-s_j|/\widetilde{\Delta t}>\sqrt{H}} |\gamma(\{t_i-s_j\}/\widetilde{\Delta t})| \left| \sum_{|t_i-s_j|/\widetilde{\Delta t}>\sqrt{H}} k_H^{'}(t_i-s_j) \right| \} = o(H^{-1}),$$

since $\sup_{|t_i-s_j|/\widetilde{\Delta t}\leq\sqrt{H}} |k_H^{'}(t_i-s_j)| \to k'(0)=0$ under the Assumption 2. By Assumption 6, $\sum_{|t_i-s_j|/\widetilde{\Delta t}\leq\sqrt{H}} \gamma(\{t_i-s_j\}/\widetilde{\Delta t}) = O(n_1 \wedge n_2)$. The last supremum term vanishes at the exponential rate by Assumption 6. The expectation of squares of (A.17) is bounded by

$$C_2(n_1\wedge n_2)^{-2}H^{-2}\left\{\sum_{i,j,r,l} EU_1(t_i)U_2(s_j)U_1(t_r)U_2(s_l)k_H^{'}(t_i-s_j)k_H^{'}(t_r-s_l)\right\}, \tag{A.18}$$

which is $O((n_1\wedge n_2)^{-1}H^{2\mu-1})$. Denote a set $\mathcal{S} := \{i,j,r,l; (t_i-t_r)/\overline{\Delta t} < H^\mu, (s_j-s_l)/\overline{\Delta s} < H^\mu\}$ where $0<\mu<1$. Then the terms in the curly bracket in (A.18) is given by

$$\begin{aligned}
&\left\{\sum_{i,j,r,l\in\mathcal{S}} + \sum_{i,j,r,l\in\mathcal{S}^c}\right\} EU_1(t_i)U_2(s_j)U_1(t_r)U_2(s_l)k_H^{'}(t_i-s_j)k_H^{'}(t_r-s_l) \\
&\leq \sup_{i,j,r,l\in\mathcal{S}} |EU_1(t_i)U_2(s_j)U_1(t_r)U_2(s_l)| \left| \sum_{i,j}\sum_{|h|,|v|<H^\mu} k_H^{'}(t_i-s_j)k_H^{'}(t_{i-h}-s_{j-v}) \right| \\
&+ C_3 n_1^2 n_2^2 \sup_{i,j,r,l\in\mathcal{S}^c} |EU_1(t_i)U_2(s_{i-h})U_1(t_r)U_2(s_{r-v})| = (i)+(ii).
\end{aligned}$$

For balanced and equally spaced case, (i) simplifies to

$$\sup_{|i-r|<H^\mu,|h-v|<H^\mu} |EU_1(t_i)U_2(s_{i-h})U_1(t_r)U_2(s_{r-v})| \left| \sum_{|i-r|<H^\mu,|h-v|<H^\mu} k_H^{'}(t_i-s_{i-h})k_H^{'}(t_r-s_{r-v}) \right|.$$

When sample size is balanced, it holds that $\frac{t_i-s_{i-h}}{\hbar} \simeq \frac{h}{H}$ under Assumption 2. Then

$$\sum_{|i-r|<H^\mu,|h-v|<H^\mu} k^{'}(\frac{h}{H})k^{'}(\frac{v}{H}) = 2H^\mu n \sum_{|h-v|<H^\mu} k^{'}(\frac{h}{H})k^{'}(\frac{v}{H}) \leq 4H^{2\mu}n\sum_{h=1}^{n}\{k^{'}(\frac{h}{H})\}^2.$$

For unbalanced case, we use the fact that $\sum_{i,j}\{k_H^{'}(t_i-s_j)\}^2 \simeq (n_1\wedge n_2)H\int_{-\infty}^{\infty}\{k'(x)\}^2dx$ and that the order of $\#\{0\leq i,r\leq n_1; |t_i-t_r|/\overline{\Delta t} < H^\mu\}$ is same as when the data is equally spaced under Assumption 2. Then $(i) = \rho(0)4(n_1\wedge n_2)H^{2\mu+1}\int_{-\infty}^{\infty}\{k'(x)\}^2dx$. We have $(ii) = C_3 n_1^2 n_2^2 \sup_{|\tau|>H^\mu}\rho(\tau)$ which is exponentially vanishing by Assumption 6. The expectation of (A.16) is given by

$$\hbar^{-2}\left\{\sum_{|t_i-s_j|/\widetilde{\Delta t}\leq\sqrt{H}} + \sum_{|t_i-s_j|/\widetilde{\Delta t}>\sqrt{H}}\right\} E\{U_1(t_i)U_2(s_j)\}e^{-is_j\alpha}k_H^{''}(t_i-s_j)\Delta t_{i+1}\Delta s_{j+1} = (i)+(ii).$$

(ii) is bounded by

$$\hbar^{-2}\sup_i(\Delta t_{i+1})\sup_j(\Delta s_{j+1})\sup_{|t_i-s_j|/\widetilde{\Delta}t>\sqrt{H}}|EU_1(t_i)U_2(s_j)|\sum_{|t_i-s_j|/\widetilde{\Delta}t>\sqrt{H}}|k_H''(t_i-s_j)|$$
$$\leq\ C_4\frac{n_1\vee n_2}{H}\sup_{|t_i-s_j|/\widetilde{\Delta}t>\sqrt{H}}\left|\gamma(|t_i-s_j|/\widetilde{\Delta}t)\right|\int_{-\infty}^{\infty}|k''(x)|dx,$$

which vanishes at the exponential rate by the Assumption 6. $\int_{-\infty}^{\infty}|k''(x)|dx$ is well defined by the Assumption 3. (i) is bounded by

$$\eta\frac{n_1\vee n_2}{H^2(n_1\wedge n_2)}\sum_{|t_i-s_j|/\widetilde{\Delta}t\leq\sqrt{H}}E\{U_1(t_i)U_2(s_j)\}k_H''(t_i-s_j)e^{-is_j\alpha}\simeq\eta\frac{n_1\vee n_2}{H^2}\left|k''(0)\right|\Gamma_{12},$$

by Assumption 6. The order of (A.16) is derived similarly as (A.18). The expectation of squares of (A.16) is bounded by

$$C_5\left(\frac{n_1\vee n_2}{H^2(n_1\wedge n_2)}\right)^2E\left\{\sum_{i=1}^{n_1-1}\sum_{j=1}^{n_2-1}U_1(t_i)U_2(s_j)k_H''(t_i-s_j)\right\}^2 \tag{A.19}$$
$$\simeq C_5\left(\frac{n_1\vee n_2}{H^2(n_1\wedge n_2)}\right)^2\rho(0)4(n_1\wedge n_2)H^{2\mu+1}\int_{-\infty}^{\infty}\{k''(x)\}^2dx=O(\frac{(n_1\vee n_2)^2}{n_1\wedge n_2}H^{2\mu-3}).$$

When data is balanced it simplifies to $O(nH^{2\mu-3})$. In summary we have

$$E\left\{\sum_{|\alpha|\leq m/2}K_H(\lambda_\alpha)\mathcal{F}(dU_1)(\alpha)\mathcal{F}(dU_2)(q-\alpha)\right\}\simeq\eta\frac{n_1\vee n_2}{H^2}|k''(0)|\Gamma_{12}$$
$$E\left\{\left|\sum_{|\alpha|\leq m/2}K_H(\lambda_\alpha)\mathcal{F}(dU_1)(\alpha)\mathcal{F}(dU_2)(q-\alpha)\right|^2\right\}=O(\frac{(n_1\vee n_2)^2}{n_1\wedge n_2}H^{2\mu-3}).$$

With some algebra it is easy to show that under the optimal bandwidth given in Theorem 2, the square root of (A.19) multiplied by the rate of convergence of the distribution $n^{\vartheta},\vartheta=\frac{2-\beta}{5}$ is $o(1)$. All other terms that involve the end terms are of smaller order by similar argument given in Lemma A.4 and Lemma A.5, Barndorff-Nielsen et al. (2011) Therefore the microstructure noise only contributes to the asymptotic bias.

A.3 Proof of Theorem 3

Our Theorem 2 implies that (8) is uniformly consistent in q. If we assume the modulus of continuity of $\mathbf{\Sigma}(t)$ is available and given by (14) then by triangular inequality, there exists sequence $\delta(n) \to 0$ such that

$$\sup_{\delta(n) \le t \le 2\pi - \delta(n)} \|\mathbf{\Sigma}(t) - \frac{1}{2\pi} \sum_{|q| \le m/2} K_H(\lambda_q) \exp(iqt) \widehat{\mathcal{F}}(\mathbf{\Sigma})(q)\|_2 \le \mathcal{C}(\frac{4}{m}).$$

Therefore we have uniform consistency result for the estimator of instantaneous covariance matrix.

REFERENCES

Aït-Sahalia, Y., Fan, J., and Xiu, D. (2010a), "High Frequency Covariance Estimates with Noisy and Asynchronous Financial Data," *Journal of the American Statistical Association*, 94, 9–51.

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2010b), "Ultra High Frequency Volatility Estimation With Dependent Microstructure Noise," *Journal of Econometrics*, 160, 160–175.

Aldous, D. J. and Eagleson, G. K. (1978), "On Mixing and Stability of Limit Theorems," *The Annals of Probability*, 6, 325–331.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001), "The Distribution of Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42–55.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008), "Designing Realized Kernels to Measure the Ex-post Variation of Equity Prices in the Presence of Noise," *Econometrica*, 76, 1481–536.

— (2011), "Multivariate Realized Kernels: Consistent Positive Semi-definite Estimators of the Covariation of Equity Prices with Noise and Non-synchronous Trading," *Journal of Econometrics*, 162, 149–169.

Barndorff-Nielsen, O. E. and Shephard, N. (2002), "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society, Series B*, 64, 253–280.

— (2004), "Econometric Analysis of Realized Covariation: High Frequency Covariance, Regression and Correlation in Financial Economics," *Econometrica*, 72, 885–925.

Christensen, K., Kinnebrock, S., and Podolskij, M. (2010), "Pre-averaging Estimators of the Ex-post Covariance Matrix in Noisy Diffusion Models with Non-synchronous Data," *Journal of Econometrics*, 159, 116–133.

Hayashi, T. and Yoshida, N. (2005), "On Covariance Estimation of Non-synchronously Observed Diffusion Processes," *Bernoulli*, 11, 359–379.

— (2008), "Asymptotic Normality of a Covariance Estimator for Non-synchronously Observed Diffusion Processes," *Annals of the Institute of Statistical Mathematics*, 60, 367–406.

Jacod, J. (1997), "On Continuous Conditional Gaussian Martingales and Stable Convergence in Law." Tech. rep., Publ. Lab. Probabilites 339, Univ. Paris-6.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., and Vetter, M. (2009), "Microstructure Noise in the Continuous Case: the Pre-averaging Approach," *Stochastic Processes and Their Applications*, 119, 2249–2276.

Jacod, J. and Protter, P. (1998), "Asymptotic Error Distributions for the Euler Method for Stochastic Differential Equations," *The Annals of Probability*, 26, 267–307.

Kalnina, I. and Linton, O. (2008), "Estimating Quadratic Variation Consistently in the Presence of Correlated Measurement Error," *Journal of Econometrics*, 147, 47–59.

Kristensen, D. (2010), "Nonparametric Filtering of the Realized Spot Volatility: a Kernel-based Approach," *Econometric Theory*, 26, 60–93.

Malliavin, P. and Mancino, M. E. (2009), "Fourier Series Method for Nonparametric Estimation of Multivariate Volatility," *The Annals of Statistics*, 37, 1983–2010.

Mykland, P. A. and Zhang, L. (2006), "ANOVA for Diffusions and Itô Processes," *The Annals of Statistics*, 34, 1931–1963.

Park, S. and Linton, O. (2011), "Realized Volatility: Theory and Application," in *Volatility models and their applications, Wiley*, eds. Bauwens, L., Hafner, C., and Sebastien, L., London, pp. 293–312.

Robinson, P. M. (1975), "Continuous Time Regression with Discrete Data," *The Annals of Statistics*, 3, 688–697.

Zhang, L. (2006), "Stochastic Volatility Using Noisy Observations: A Multi-Scale Approach," *Bernoulli*, 12, 1019–1043.

— (2010), "Estimating Covariation: Epps effect, Microstructure Noise," *Journal of Econometrics*, 160, 33–47.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005), "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100, 1394–1411.